Inclusive environmental governance *pp. 1272 & 1274*

Using donor T cells for cancer immunotherapy *pp. 1275 & 1337*

Basalts can rapidly trap carbon dioxide *p. 1312*

Science

$15
10 JUNE 2016
sciencemag.org
AAAS

RED ALERT

The fight to save Britain's red squirrels *p. 1268*

Science

AAAS

1200 New York Avenue, NW, Washington, DC 20005
Bateman House, 82-88 Hills Road, Cambridge, UK CB2 1LQ

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

INSTITUTIONAL SITE LICENSES 202-326-6730 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Services** 202-326-6417 or http://membercentral.aaas.org/discounts

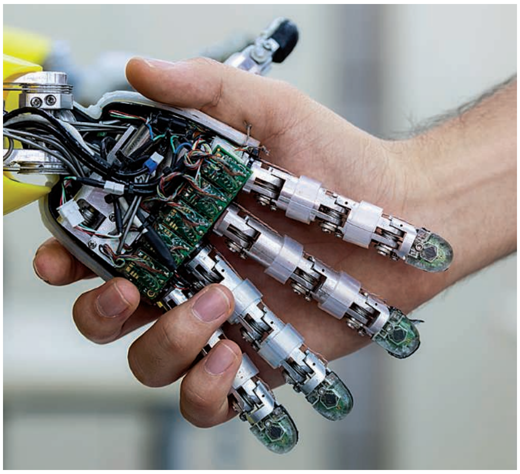
Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority of conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

INFORMATION FOR AUTHORS See pages 624 and 625 of the 5 February 2016 issue or access www.sciencemag.org/authors/science-information-authors

EDITORIAL

Robotics takes off

In a mere 50 years, robots have gone from being a topic of science fiction to becoming an integral part of modern society. They now are ubiquitous on factory floors, build complex deep-sea installations, explore icy worlds beyond the reach of humans, and assist in precision surgeries. These robots are improving our health, adding to our scientific understanding, raising productivity, and performing tasks that humans cannot do. The rapid increase in the capabilities, varieties, and applications of robots has been built on scientific and engineering research into power and actuation systems, artificial intelligence, onboard navigation, environmental sensors, manipulators, control systems, novel materials, microfluidics, systems integration, and many other advances. With this growth, the research community that is engaged in robotics has expanded globally. To help meet the need to communicate discoveries across all domains of robotics research, we are proud to announce that *Science Robotics* is open for submissions.



"Robots...will increasingly be intertwined with us..."

One measure of the growth in the field of robotics is the increasing number and size of robotics conferences. Over a dozen prominent general robotics conferences are held annually at venues around the globe, with additional domain-specific conferences being convened as well. Online conference proceedings have traditionally served the role of journals for widely disseminating the material presented at conferences, giving credit to authors for their work, and archiving results. However, as the field has matured, publishers have launched new journals to more broadly communicate results to all interested scientists, regardless of their ability to attend conferences.

Science Robotics aims to select the most groundbreaking advances in robotics across applications (such as medical, industrial, land, sea, air, space, and service), systems (such as propulsion, sensors, control, and navigation), and scales (from macro to nano) of general interest to the robotics research community and researchers working in allied fields (such as bio-inspired engineering, materials science, and novel sensing technologies). The goal is to move the field forward and cross-fertilize different research applications and domains. The *Science Robotics* Editorial Board will screen submissions for the most original research and then apply *Science*'s rigorous peer-review process to ensure that the papers published are well worth reading. The journal will also publish invited reviews and will develop a forum to explore current policy, ethical, and social issues that affect the robotics community.

We aim to make *Science Robotics* the must-read journal for the latest discoveries in robotics that will drive the next generation of robots. To this end, our editors welcome papers on relevant advances in other disciplines with a strong potential to revolutionize the design or operation of robots. Eventually, we plan to engage with robotics educational programs, using *Science Robotics* content.

Authors are still encouraged to submit to *Science* outstanding robotics papers of interest to communities well beyond robotics researchers. We have a simple process for transferring robotics papers among the various relevant journals in the *Science* family (*Science*, *Science Advances*, *Science Robotics*, and for medical robotics, *Science Translational Medicine*) that permits the reuse of reviews if reviewers agree, speeding up the process of a paper being sequentially considered at multiple journals.

Robotics is still a relatively young field. The problems prompting society to develop robotic systems to replace or extend human presence are unlikely to abate. For example, many nations face an aging population, with an insufficient number of young people to both care for the elderly and support the needs of the economy. Robots can accomplish some routine, dangerous, or high-precision jobs better than humans, thereby freeing people to do the tasks that robots cannot do. Robots are already part of our modern society and will increasingly be intertwined with us economically, technically, and perhaps, socially. *Science Robotics* seeks to help shape the future scientific, technical, ethical, and social aspects of that evolution.

– Guang-Zhong Yang and Marcia McNutt

Guang-Zhong Yang is the editor of Science Robotics *and director of the Hamlyn Centre, Imperial College London, UK. Email: g.z.yang@imperial.ac.uk*

Marcia McNutt Editor-in-Chief Science *Journals*

PHOTOS: (INSET) MURATSENEL/ISTOCKPHOTO.COM; (TOP RIGHT) IMPERIAL COLLEGE LONDON; (BOTTOM RIGHT) STACEY PENTLAND PHOTOGRAPHY

10.1126/science.aag2724

Published by AAAS

NEWS

"Neuroscience is in its infancy—it's where astronomy was under Galileo. It's a wide open horizon."

Psychologist and writer Maria Konnikova, speaking 4 June at the World Science Festival in New York City.

IN BRIEF

Extinct steppe bison fossils helped researchers determine when a corridor opened between Canada's ice age glaciers.

Humans didn't wait for melting ice to settle Americas

For most of the last ice age, enormous glaciers covered western Canada—yet people managed to cross into the Americas from Alaska. Now, a new study of bison fossils may answer a longstanding question: How did the earliest migrants make the journey? Many researchers thought that Clovis hunters traveled through a narrow land passage between glaciers more than 13,000 years ago, when their distinctive fluted spear points appeared in the New World. But there have been little hard data on when such a passage opened. So scientists turned to a new tool: bison fossils. The growth of the glaciers had separated bison populations in North America into northern and southern branches. Those populations, the team found, were distinguishable by their mitochondrial DNA (passed down from the mother); finding a northern bison in the south (or vice versa) would point to an ice-free passage. After sequencing the fossils' mitochondrial DNA, the team dated them using radiocarbon in their collagen. The bison were mingling in the ice-free corridor by about 13,000 years ago, the team reports online this week in the *Proceedings of the National Academy of Sciences*. That's right around the first appearance of Clovis points—but doesn't account for a handful of pre-Clovis sites. Those earliest migrants, the researchers say, probably hopped down the Pacific coast in boats. http://scim.ag/bisonicepassage

AROUND THE WORLD

A human genome from scratch?

BOSTON | A group of researchers is floating a proposal for a 10-year project to assemble a human genome from its chemical components. The Human Genome Project–Write (HGP-write) would aim to lower the cost of engineering large stretches of DNA and testing their activity in cells. That could lead to safer disease treatments or new organs for human transplant, the group says. An invitation-only meeting about the project at Harvard Medical School in Boston last month fueled fears that scientists might use engineered cells to create designer humans and led some researchers to call for a broader public discussion. The proposal, published online in *Science* on 2 June, suggests an initial $10 million in funding to launch

PHOTO: GOVERNMENT OF YUKON

HGP-write in 2016, and promises "open and ongoing dialogue" about the work's ethical, legal, and social implications. http://scim.ag/humangenomeprop

U.K. loses too much research

LONDON | Government agencies in the United Kingdom do a poor job of keeping tabs on research the agencies fund to set policies, according to a report released 1 June by Sense about Science, a London-based group that advocates for the use of scientific evidence in policymaking. The group queried 24 government agencies to find out how they publish and archive research that they commission. Only four departments maintained a database of their research; 11 could not create an inventory of what they had supported, mostly because no central records exist. Agency confusion over when research reports should be published has resulted in what the group calls "ghost research" that is "unsearchable in the national archives and exists only in the memories of officials." The report also described examples of delays in releasing the results of what it called "politically awkward" studies. http://scim.ag/UKtrackpolres

Gene drive gets a yellow light

WASHINGTON, D.C. | It may take 5 years or more before researchers will be ready to try a controversial technology that eradicates or alters disease vectors by rapidly driving a genetic change through their populations. Nevertheless, researchers, funding organizations, and regulatory agencies should hasten to grapple with societal and regulatory issues stemming from this "gene drive" technology, the National Academies of Sciences, Engineering, and Medicine urged in a report released this week. The report notes that the technology offers great promise for agriculture, conservation, and public health; if applied to mosquitoes, for example, it could rid an area of malaria-carrying vectors. But, it stresses, the current regulatory system—which includes institutional review boards and environmental impact assessments—is not adequate to address the potentially great risks posed by gene drive–altered organisms. It calls for much greater involvement of the public in the early stages of the technology's development and approval for use, and emphasizes the need to bring developing countries up to speed, as that is where such technology will likely be tried first. http://scim.ag/_genedrive

Mosquito-borne disease threatens this honeycreeper.

Bacterial activity in seafloor soil produced this "column."

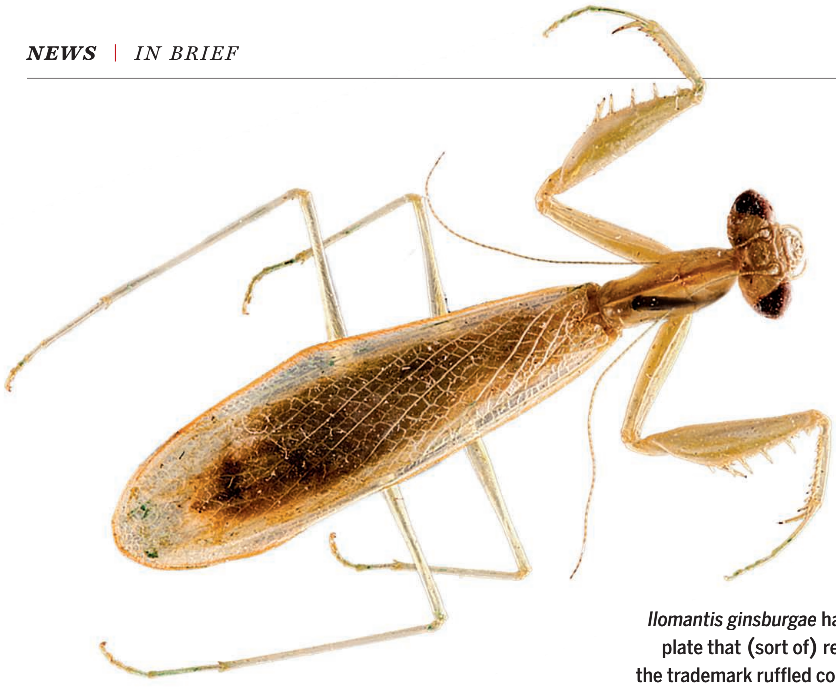
'Lost city' was made by microbes

In 2014, tourists diving off the Greek island of Zakynthos snapped some photos of several mysterious structures on the sea floor—resembling the remnants of a paved stone walkway and colonnades—and uploaded them to Google Earth. But had they really found a long-lost city? Greece's Ephorate of Underwater Antiquities investigated but found no human artifacts such as coins or pottery. So a team of scientists, including geochemist Julian Andrews of the University of East Anglia in Norwich, U.K., stepped in to study the structures. The team examined the underwater terrain, as well as the isotopic values of carbon, oxygen, and strontium in the structures—and determined that microbes, rather than humans, had "built" them, they reported online last week in *Marine and Petroleum Geology*. Sulfate-reducing bacteria in the sediments, munching on methane seeping up through the sea floor, can create favorable conditions for dolomite, a type of magnesium-rich carbonate rock, to form within the soil. Diffuse methane seeps result in sheetlike structures similar to paving stones. More focused seeps lead to round structures resembling the bases of columns. The team's multilayered approach to solving the mystery is "a case study on tracking down microbes and their involvement," Andrews says. "It lays out the way that you would unfold the detective story."

LISA Pathfinder passes test

MADRID | The European Space Agency (ESA) launched LISA Pathfinder in December 2015 to see whether it is possible to build a gravitational wave detector in space. The answer? A resounding yes, mission operators announced at a press conference 7 June. For 55 days, the space-based detector used a laser interferometer to continuously measure the distance between two free-floating metal cubes—"test masses" influenced solely by gravity—within the spacecraft. The goal was to assess whether the cubes can be truly isolated from other forces, including solar wind and radiation. If so, ESA has proposed the €1 billion Evolved Laser Interferometer Space Antenna (eLISA), consisting of three spacecraft flying in formation millions of kilometers apart, and firing laser beams between them to measure the distances. If a gravitational wave (generated by the merging of black holes or an exploding supernova) passes by, it will compress and stretch space, causing those distances to change. The changes will be tiny—perhaps a millionth of a millionth of a meter—so the

PHOTOS: (LEFT TO RIGHT) © PHOTO RESOURCE HAWAII/ALAMY STOCK PHOTO; THE EPHORATE OF UNDERWATER ANTIQUITIES



Ilomantis ginsburgae has a neck plate that (sort of) resembles the trademark ruffled collar of its namesake, Ruth Bader Ginsburg.

Supreme Court justice inspires mantis name, insect classification

In a break from the traditional use of male genitalia to classify insect species, scientists have for the first time formally used female genitalia to identify a new species of praying mantis. Male genitalia have long been preferred because of their seemingly wider—and more easily observed—variety of shapes and sizes. For example, insects with a hook-shaped penis might be classified as one species, whereas those with slightly curved genitalia might be classified as another. However, the researchers found that they could use female genital features alone to define *Ilomantis ginsburgae*, a leaf-dwelling mantis from Madagascar, as they reported last week in *Insect Systematics & Evolution*. The mantis was named in honor of U.S. Supreme Court Justice Ruth Bader Ginsburg, a strong supporter of gender equality and a regular wearer of jabot collars, which resemble the neck plate of the insect.

measurement systems must be precise. And so far, the mission is meeting eLISA requirements, ESA reported at the conference and in a paper published this week in *Physical Review Letters*.

3D-printed stations come online

LUSAKA | Five low-cost weather stations constructed primarily out of 3D-printed, easily replaceable parts have come online in different locations around Zambia, providing forecasts to local farmers to help them optimize planting and harvesting times and minimize impacts from natural hazards such as floods. The project is a collaboration between the National Center for Atmospheric Research in Boulder, Colorado, and the Zambia Meteorological Department in Lusaka. After testing the stations last year, the team installed one of the stations next to the meteorological department, another three next to local radio stations, and the fifth near a rural hospital. Zambia is to take over management of the project by the end of 2016; the goal is to build a network of 100 3D-printed weather stations in the country, at a cost of about $300 per station.

Poor review leads to NIH uproar

BETHESDA, MARYLAND | A decision to overhaul the leadership of the National Institutes of Health (NIH) Clinical Center after an outside review group found serious patient safety problems has sparked an uproar. In a recent letter, department chiefs at the center wrote that the review, triggered by problems with a drug production facility, unfairly concluded that patient safety has been compromised across the research hospital (*Science*, 20 May, p. 875). They say the working group's report has demoralized staff, worried patients, and "demonized" the center's leadership. Patient advocates and clinical research leaders across NIH have also written letters taking issue with the review. In a statement on 2 June, NIH Director Francis Collins responded to a letter from Clinical Center department heads, noting that he is taking the comments "very seriously," but "stand[s] by" the outside working group's process and expertise and agrees that the center needs "more central authority and accountability." http://scim.ag/NIHuproar

BY THE NUMBERS

$2 billion

The increase, to $34.1 billion, that a U.S. Senate appropriations panel wants to give the National Institutes of Health. But Congress likely won't settle on a 2017 budget number before November.

£25 million

New infusion of money for the open-access journal *eLife*, created in 2012 by three biomedical giants. The funders are providing the additional funds for 2017–2022.

NEWSMAKERS

Bird flu scientist flies out

Italian virologist-turned-politician **Ilaria Capua**, 50, is leaving Italy and returning to science. Capua, an avian influenza expert, has given up the seat in the Italian Chamber of Deputies that she won in 2013 to become head of the One Health Center of Excellence for Research and Training at the University of Florida in Gainesville later this month. Capua is the former director of the Division of Biomedical Science of the Istituto Zooprofilattico Sperimentale delle Venezie, a government lab for veterinary science; she became known for her advocacy for public access to genetic data about influenza. Along with 40 others, Capua is a suspect in a judicial investigation into virus smuggling, spreading bird flu, and other alleged crimes (*Science*, 5 September 2014, p. 1105). The investigation, based solely on tapped phone calls, is ongoing; Capua denies any wrongdoing. Her stint in parliament has been frustrating, she says: "Politics is a complicated world, especially if you think in a rational and fact-related fashion." http://scim.ag/CapuaUF

PHOTO: RICK WHERLEY/CLEVELAND MUSEUM OF NATURAL HISTORY

The hobbit, shown in a reconstruction, may have had a long history on the Indonesian island of Flores.

HUMAN ORIGINS

Likely hobbit ancestors lived 600,000 years earlier

Fragmentary "hobbitlike" fossils show evolution of dwarf human on Indonesian island

By **Elizabeth Culotta**

From the moment that the announcement of a 1-meter-tall ancient human nicknamed "the hobbit" shocked the world in 2004, supporters and skeptics alike have longed for more fossils. After the first couple years of discoveries, the research team kept digging, hoping to shore up the creature's status as a separate species and settle the mystery of its origins. They dug at the original find site, they dug elsewhere on the Indonesian island of Flores, and they dug on nearby Sulawesi. They unearthed thousands of stone tools and tens of thousands of animal bones. But they found no human fossils. Until now.

This week in *Nature*, the team announces that they have found specimens of a tiny hominin at a site on Flores called Mata Menge, 74 kilometers from the hobbit's home in Liang Bua cave. The haul is meager—a fragment of jaw and isolated teeth—but the fossils' diminutive size suggests they belong to the hobbit's species, *Homo floresiensis*, or a precursor to it. They are securely dated to 700,000 years ago, hundreds of thousands of years earlier than the hobbit—and they are about 20% smaller. Their size is "amazing!" says Christoph Zollikofer of the University of Zurich in Switzerland, who studies fossils of the human ancestor *H. erectus* from Dmanisi, Georgia.

To many, the finds suggest that a lineage of tiny humans evolved on Flores, emerging surprisingly soon after their ancestor, likely *H. erectus*, arrived about 1 million years ago. "We expected to find a large-bodied, close relative of *Homo erectus*," says paleontologist Gerrit van den Bergh of the University of Wollongong in Australia, co-leader of the discovery team. "Instead we found fossils of tiny humans, even slightly smaller than *H. floresiensis*!"

The scanty new fossils are the fruit of a herculean effort at Mata Menge, in the Soa Basin on Flores. About 50 years ago, a prescient Dutch Catholic priest and amateur archaeologist discovered stone tools there and concluded that *H. erectus* had washed up on Flores, perhaps from nearby Java. No one believed him at the time, but researchers, including members of the current team, have scoured the basin ever since. Grassy vegetation and a humid climate obscure and destroy any fossils that peek from the surface, so the team had to dig through swaths of the landscape wholesale. They bulldozed off top layers of soil, employing up to 120 students and local workers to sieve tons of dirt and chisel out fossils over 10 field seasons.

Finally, in the last few weeks of the project's final season in 2014, their labors paid off. A sharp-eyed worker spotted a hominin tooth, which was followed by a bit of lower jaw and cranium plus five more teeth, including two baby teeth.

The jaw—clearly that of an adult, as a wisdom tooth had erupted—is about 20% smaller than the two recovered from Liang Bua, which are about the size of a 5-year-old modern human child's. The relatively thin body of the jaw and a crest on one molar link the fossils to *H. erectus* and *H. floresiensis*, but not to the earlier human ancestor *Australopithecus*, the team reports. That argues against one scenario for hobbit origins—that a very small, primitive hominin such as *Australopithecus* somehow got out of Africa and to Flores (*Science*, 14 October 2005, p. 208).

Because the fossils are fragmentary, "we want to have other skeletal elements before we finally conclude about taxon," says team member Yousuke Kaifu of the National Museum of Nature and Science in Ibaraki, Japan. For now, he and his colleagues simply call it "*Homo floresiensis*–like."

To arrive at the 700,000-year figure for the

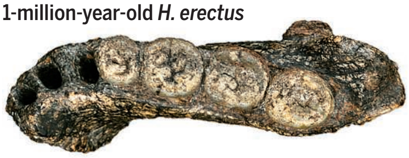
A dwarf lineage

The newly discovered jaw is the smallest yet: It is smaller than that of the original hobbit and much smaller than that of *Homo erectus* from Java.

60,000-year-old hobbit

700,000-year-old likely hobbit ancestor

1-million-year-old *H. erectus*



PHOTOS: (LEFT TO RIGHT) © B. CHRISTOPHER/ALAMY STOCK PHOTO; Y. KAIFU *ET AL.*/*NATURE* © JUNE 2016

bones' age, the team used radiometric techniques to date volcanic layers above and below the soil layer where they were found, and also directly dated a partial hominin tooth. From abundant animal and plant remains they built a picture of the ancient environment: savannalike grasslands watered by meandering streams and populated by pygmy elephants, giant rats, freshwater crocodiles, and carnivorous Komodo dragons.

They also examined 149 simple stone tools uncovered near the hominins, which are mostly similar to thousands uncovered elsewhere on Flores, says team member Mark Moore of the University of New England in Armidale, Australia. He notes that one kind of relatively complex tool characteristic of *H. erectus* appears in the Flores record only about 1 million years ago, and vanishes after that. But the other, simple tools, perhaps made by hobbitlike people, remain unchanged for hundreds of thousands of years.

Putting all the evidence together, Van den Bergh and project co-director Adam Brumm of Griffith University, Nathan, in Brisbane, Australia, tell hobbit history this way: A little more than 1 million years ago, a small group of *H. erectus* was marooned on Flores, perhaps washed in by tsunamis from the islands of Java or Sulawesi. Because food was scarce and small bodies were adaptive, they quickly evolved to be smaller in size, a process called island dwarfing that is known to affect other vertebrates and also produced the island's pygmy elephants.

Several other researchers agree with this picture and praise the team's excavation. The finds "end the argument" that *H. floresiensis* is a diseased modern human rather than a separate species, as some critics have argued, says paleoanthropologist Russell Ciochon of the University of Iowa in Iowa City, who was not part of the team. (That view was already tottering after the team recently redated the original hobbit skeleton to at least 60,000 years ago, thousands of years before modern humans apparently reached the region.)

Still, the fragmentary nature of the fossils leaves parts of the story open to interpretation. "The authors have done a good job with what they have—but they don't have a lot," says paleoanthropologist Susan Antón of New York University in New York City. Zollikofer says it's possible that the Mata Menge and Liang Bua remains reflect separate colonizations and dwarfing events, given the 600,000 years that separate them. And skeptic Robert D. Martin of the Field Museum in Chicago, Illinois, says he won't be convinced that the hobbit's puny brain isn't pathological until a second skull emerges.

For the team, the road ahead is clear. Van den Bergh is succinct: "More digging." ■

MICROBIOME RESEARCH

Banking on stool despite an uncertain future

Regulation and commercialization of fecal transplants could make poop providers obsolete

By **Tina Amirtha**, *in Leiden, the Netherlands*

Since February, five volunteers from this ancient university town have been dropping by the academic hospital several times a week with a gift that is both precious and a bit distasteful: a blue plastic container holding fresh stool. Two technicians process each donation into a suspension resembling chocolate milk, then store it at –80°C until it's needed.

The volunteers are donors at the new Netherlands Donor Feces Bank (NDFB), which last month began shipping its stool preparations to hospitals around the Netherlands for use in fecal microbiota transplants (FMTs), a procedure in which doctors seek to restore the normal microbial balance in a patient's gut with a dose of stool from a healthy volunteer.

Similar banks have opened their doors in the United States, France, and the United Kingdom. In Germany, a group of physicians is seeking permission to operate a national stool bank at the University of Cologne. And groups elsewhere in Europe, Latin America, and Asia are also interested in starting banks, says Carolyn Edelstein, director of policy and global partnerships at OpenBiome, a stool bank based in Medford, Massachusetts.

Yet the long-term future for these efforts is unclear. The banks, all nonprofits, help meet a growing demand for safe stool preparations at a reasonable price—the Leiden bank charges €145 per dose—and aid research into FMTs. But regulators on both sides of the Atlantic have yet to decide on how to regulate the banks' products, and several companies are busy developing gut microbiome replacements that lack the yuck factor and could, in the long run, make the stool banks obsolete.

An FMT can be administered as an enema or through a tube that goes down the nose and into the stomach. So far, the procedure has been shown to aid only one condition in a randomized clinical trial: A landmark Dutch study published in 2013 showed that it prevented relapse in patients with recurrent infections of the intestinal bacterium *Clostridium difficile*. But many scientists believe that by restoring microbial harmony the treatment could help other diseases associated with an abnormal gut microbiome, such as irritable bowel syndrome, Crohn's disease, and ulcerative colitis. Several companies have invested heavily in the concept, and at least 100 clinical trials are planned or underway, 16 of them in China alone.

Banks on the rise

Five stool banks have opened since 2012; many more are planned.

2012, OpenBiome (Medford, Massachusetts)

2014, University Hospitals of Paris Centre

2015, AdvancingBio (Mather, California); **Public Health England, Birmingham laboratory** (United Kingdom)

2016, Netherlands Donor Feces Bank (Leiden, the Netherlands)

Many physicians and hospitals concoct the infusions themselves, often using donations from a patient's family member and sometimes with store-bought kitchen blenders. Although a few ground rules have developed—fresh stool is mixed with saline solution and filtered—details and safety precautions vary widely. It's also a messy and time-consuming business.

That's where the banks come in. OpenBiome, launched in 2012, has already provided stool for 13,000 FMTs in the United States and in six other countries; now, others are following suit (see box, above). The banks have all developed their own procedures and safeguards, often learning from each other. NDFB, launched by the authors of the 2013 study, spent more than a year drawing up protocols for collecting, processing, and storing stool, says co-founder Ed Kuijper of Leiden University Medical Center.

To ensure good microbial diversity and a minimum of pathogens, criteria for donors are strict; at the Leiden bank, the current five donors were the only ones deemed eligible from among more than 200 candidates. Being obese or above 50 is disqualifying, for instance, as are risky sexual behaviors and

recent travel to countries where intestinal pathogens are rife. (Smoking or using marijuana, on the other hand, is no problem.) After filling out a questionnaire, about one in 10 applicants is asked to submit a sample, which is screened for more than 50 potential pathogens. OpenBiome has similarly strict criteria; only 2.8% of applicants make the cut. (OpenBiome pays its volunteers $40 per donation; Dutch law bans such payments.)

Regulatory agencies have yet to catch up. The U.S. Food and Drug Administration (FDA) has decided to treat fecal transplants as it would a biological drug, requiring doctors to file a so-called Investigational New Drug (IND) application when they want to administer an FMT. The agency has waived this requirement for *C. difficile*—but new draft guidelines, released in March, limit the exemption to hospitalmade preparations that use stool from a known donor.

Stool from banks would not fall under the exemption because FDA sees it as more risky. One reason is that relatively few, anonymous donors provide stool for many patients, meaning that any pathogens a donor harbors could spread widely. If the guidelines are adopted, U.S. hospitals might stop using stool banks, a prospect that has alarmed patients and FMT advocates. They worry that access to the procedure will become harder and dispute that hospitals' own products are safer.

In Europe, the regulatory future is unclear as well. No rules for FMTs exist at the E.U. level. Some countries, including the United Kingdom, France, and Germany, regulate FMTs as drugs, as FDA does; others have no specific regulation at all.

Meanwhile, several companies are developing new FMT products that could put the banks out of business. Rebiotix in Roseville, Minnesota, makes an FMT preparation using its own stool donors that comes with a guarantee that each suspension contains a minimum number of bacteria of sufficient diversity. Rebiotix finished a phase II trial for recurrent *C. difficile* in January, the results of which have yet to be published.

Other companies, such as Vedanta Biosciences in Cambridge, Massachusetts, hope to move away from stool altogether by growing specific bacterial strains in the lab. "It's sort of a natural progression, just like for aspirin, which started off with willow bark, and then we figured out that you could actually just synthesize the active component," says OpenBiome's co-founder and research director Mark Smith.

If these products win FDA's approval, Edelstein says, "clinicians will have to choose between a licensed biologic product or a stool bank. It might make it harder for [them] to work with us." Smith says that the higher prices that companies will need to charge to recoup their investments could prevent some patients from obtaining treatment; he says stool banks should remain a low-cost alternative.

Whatever the commercial future of FMTs, the stool banks say they'll have other work to do. OpenBiome may focus more on research, Edelstein says. Besides stool, OpenBiome provides guidance on experimental designs, safety protocols, and IND applications. The Leiden bank seeks to advance science as well. It just started a research collaboration with Vedanta Biosciences, and it plans to study whether fecal transplants should be given to *C. difficile* patients at an earlier stage. "Now, patients receive [an] FMT when they have tried all the other options," Kuijper says. "But more can easily benefit." ■

Tina Amirtha is a freelance writer in Leiden, the Netherlands.

The Netherlands Donor Feces Bank has accepted only five stool donors so far.

PHOTO: MICHEL GROEN

GEOCHEMISTRY

New solution to carbon pollution?

Instead of sending CO_2 up a smokestack, researchers in Iceland turn it into rock

By **Eli Kintisch**

Researchers working in Iceland say they have discovered a new way to trap the greenhouse gas carbon dioxide (CO_2) deep underground: by changing it into rock. Results published this week in *Science* (see p. 1312) show that injecting CO_2 into volcanic rocks triggers a reaction that rapidly forms new carbonate minerals—potentially locking up the gas forever. The technique has to clear some high hurdles to become commercially viable. But scientists say the project, dubbed CarbFix, offers a ray of hope for beleaguered efforts to fight climate change by capturing and storing CO_2 from power plants. "This is a great step forward," says Sally Benson of Stanford University in Palo Alto, California, a geologist unaffiliated with the project.

Dozens of pilot projects around the world have sought to test carbon capture and storage (CCS) as a way of curbing CO_2 emissions from power plants. Very few have been scaled up, owing to prohibitive costs, estimated at $50 to $100 per ton of CO_2 sequestered.

CCS also faces technical hurdles, and one of the largest is where to store the captured gas. Most researchers favor formations of sedimentary rock, often sandstone harboring briny groundwater or depleted oil wells, because industry has long experience in working with them. But scientists fear that fissures in the rock layers that cap the storage aquifers could let CO_2 leak back into the atmosphere.

So in 2006, Icelandic, U.S., and French scientists proposed a different approach: injecting CO_2 into underground layers of basalt, the dark igneous rock that underlies Earth's oceans and crops up in parts of continents as well. They knew that unlike sandstone, the basalt contains metals that react with CO_2, forming carbonate minerals such as calcite—a process known as carbonation. But they thought the process might take many years.

Carbon dioxide pumped into deep wells in Iceland underwent a surprising chemical transformation.

PHOTO: J. MATTER

To find out, they launched the CarbFix experiment 25 kilometers east of Reykjavik, intending to dose Iceland's abundant underground basalt with CO_2 that bubbles from cooling magma underground and is collected at a nearby geothermal power plant.

In 2012, the researchers injected 220 tons of CO_2—spiked with heavy carbon for monitoring—into layers of basalt between 400 and 800 meters below the surface. They also added extra water, which reacted with the gas to form a key driver of mineral reactions, carbonic acid. Then they monitored the pH, geochemistry, and other characteristics of the subsurface by taking samples from nearby wells.

What happened next startled the team. After about a year and a half, the pump inside a monitoring well kept breaking down. Frustrated, engineers hauled up the pump and found that it was coated with white and green scale. Tests identified it as calcite, bearing the heavy carbon tracer that marked it as a product of carbonation.

Measurements of dissolved carbon in the groundwater suggested that more than 95% of the injected carbon had already been converted into calcite and other minerals. "It was a huge surprise that the carbonation happened so fast," says Juerg Matter, a geologist with CarbFix at the University of Southampton in the United Kingdom. Laboratory tests by Matter's team and others, along with computer modeling, had previously suggested that carbonation in basalt would take at least a decade. (Sandstone aquifers are so unreactive that carbonation is thought to take centuries at conventional CCS sites.)

The speedy carbonation "means this method could be a viable way to store CO_2 underground—permanently, and without risk of leakage," Matter says. Unpublished data from a similar project in basalt near the Columbia River near Wallula, Washington, point to a similar conclusion. And there is no lack of basalt formations on land or offshore, which could make CCS possible for power plants "not near sedimentary rocks or depleted oil wells," Matter adds.

Bigger field tests are needed, says geologist Peter Kelemen of Columbia University, to confirm that such a high fraction of the injected carbon was mineralized. (Columbia is a CarbFix partner, but Kelemen is not on the project.) Scaled-up demonstrations could also make sure that the speed of the reaction won't turn into a drawback, Stanford's Benson says. If carbonation generates minerals that quickly plug the pores in the basalt, she worries, they could trap CO_2 near the injection site instead of letting it spread through the rock.

But even CarbFix's own scientists acknowledge that the biggest obstacle to CCS in basalt is financial: Power companies have little incentive to pursue it. "Without a price on carbon emissions, there's no business case," admits Matter, who hopes policymakers will create such an incentive. Otherwise, projects in basalt could suffer the same fate as the dozens of conventional CCS projects around the world that have failed to be commercialized. In the meantime, says Benson, the success in Iceland is a welcome development. "We could all use some positive news in this field," she says. ■

PSYCHOLOGY

Mechanical Turk upends social sciences

Growing pains arise for researchers using online platform

By **John Bohannon**

In May, 23,000 people voluntarily took part in thousands of social science experiments without ever visiting a lab. All they did was log on to Amazon Mechanical Turk (MTurk), an online crowdsourcing service run by the Seattle, Washington–based company better known for its massive internet-based retail business. Those research subjects completed 230,000 tasks on their computers in 3.3 million minutes—more than 6 years of effort in total.

The prodigious output demonstrates the popularity of an online platform that scientists had only begun to exploit 5 years ago (*Science*, 21 October 2011, p. 307). In 2011, according to Google Scholar, just 61 studies using MTurk were published; last year the number topped 1200. "This is a revolution in social and behavioral science," says psychologist Leib Litman of the Lander College for Men in New York City, who generated the May data from TurkPrime, a website that he created last year with computer scientist Jonathan Robinson, also at Lander, to facilitate MTurk studies. "Research is moving from the lab to the cloud."

Why bother with the cloud? A social sciences study with hundreds of live subjects normally requires weeks of work just to gather the data, not to mention finding people and signing them up. Last month's studies on MTurk—which include a test of the limits of people's generosity, a comparison of religiosity and humility, and a measurement of the psychological impact of graphic warnings on cigarette packages—took only days from start to finish.

But the platform's popularity has raised concerns, as researchers discussed at the Association for Psychological Science meeting in Chicago, Illinois, last month. Some worry that they are becoming too dependent on a commercial platform. "Academic research would be really screwed if Amazon decided to shut it down," says Todd Gureckis, a psychologist at New York University (NYU) in

New York City. Others question whether the research volunteers are paid fairly and treated ethically. And looming over it all are questions about who these anonymous volunteers actually are, and concerns that they are less numerous and diverse than researchers hope.

MTurk's ascendancy in the social sciences—more than 1000 researchers have registered experiments using it on TurkPrime—is unexpected given the clunkiness of its interface. When Litman first tried to use the platform, he found it baffling. "It looks like a website designed in the 1990s by computer engineers," he says. That shouldn't be surprising, considering that Amazon created MTurk as a tool for harnessing humans to improve artificial intelligence software. For example, when a computer struggles to identify the content of a photograph, Turkers can be hired to name objects, helping the computer learn.

But researchers have more complex needs, and adapting MTurk for social sciences often requires computer programming skills that few have. Besides TurkPrime, another research tool for MTurk is psiTurk, created by Gureckis, which is like an "app store" for experiments. Rather than write programs from scratch, you can browse free, open-source code from other researchers' experiments and just tweak it.

The thorniest issue for researchers has been getting a handle on just how big and diverse the Turker population is. Amazon has boasted that MTurk harnesses more than 500,000 workers from around the globe, but what researchers want to know is how many unique, active users are willing to participate in their studies at any given time. If that number is small, then the same people could be recirculating through experiments, and that can bias the results.

When Turkers register, Amazon marks them with an ID that is buried in the raw code after a work session is complete. Those IDs have enabled researchers to study Turker demographics with a method borrowed from wildlife ecology called capture-recapture: To estimate the number of fish in a lake, capture them, mark them, and return them; the smaller and slower changing the population, the higher the proportion of marked fish will get recaptured. By comparing the Turker IDs from experiments across multiple labs, it is possible to conduct a virtual capture-recapture survey retrospectively.

Neil Stewart, a psychologist at the University of Warwick in Coventry, U.K., led the first effort to estimate the effective MTurk research population with this method—and the results sent shock waves through the community last year. Seven psychology labs in the United States, Europe, and Australia ran 114,000 experimental sessions over a 3-year period. The number of unique people among the subjects came to only 30,000. Rather than a pool of half-a-million subjects always on tap, Stewart estimated that the true number of Turkers that are willing to take part in an experiment at any one time is only about 7300.

"What seemed like a virtually infinite subject pool was in fact more like a very large state university psychology pool," Gureckis says. Stewart's data show that the population churns rapidly: Half the Turker population that participates in research is replaced by fresh people every 7 months.

Those Turkers are also far less diverse than was thought. Though Amazon has long noted the global nature of the community, surveys of those completing experimental tasks reveal that the vast majority are based in the Unites States. And compared with the average American, Litman says, Turkers "skew young, they are more liberal, more urban, and more likely to be single." Knowing such traits, he notes, is crucial for researchers as they try to interpret their data.

Turkers are also poorly paid, although their hourly rate is difficult to calculate, in part because Amazon takes a cut of between 20% to 40%, and because Turkers undertake multiple tasks at once with break time. For the 6 years of accumulated effort in May, researchers doled out $164,882. That would seem to translate to an average pay rate of about $3 per hour, but "the true hourly rate is somewhere between $4 and $8 per hour," Litman says.

Many Turkers complain that this is too low, because social science experiments often take more effort—and time—than other tasks. Gureckis agrees. At NYU, "we pay subjects $8 to $10 per hour, and there's often a bonus at the end," he says. "MTurk subjects should be paid the same as they would in the lab. That's what we try to do."

Nor do they yet enjoy the same ethical protections. When subjects drop out of an experiment, "you're supposed to pay them proportional to the time they put in," Gureckis says. But MTurk has no mechanism for partial pay: A Turker must complete an entire task or get no pay at all.

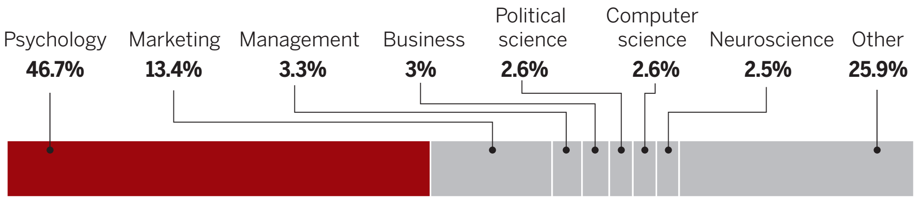
And although researchers are supposed to protect the identity of subjects, Gureckis says, "MTurk is not really anonymous." In 2013, a research team showed that it is possible to match a Turker's worker ID to their account on Amazon's retail website. Depending on how much information is associated with that user profile, it can reveal a Turker's buying habits, video tastes, and even their full name and location. Researchers have suggested ways to remedy this privacy issue, but the company so far seems to have taken no action. (Amazon declined a request for an interview with *Science*.)

Some researchers wonder whether the giant company will keep MTurk going—or whether social scientists need to develop their own customized alternative. "At this point, MTurk has become so important for social science that the National Science Foundation should be negotiating directly with Amazon," Gureckis says. "We're subsidizing this service with millions of dollars in federal grant money." ■

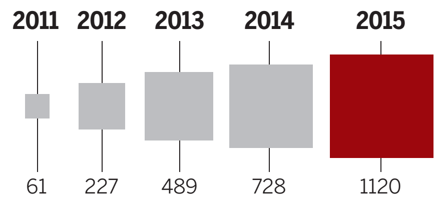
Amazon Mechanical Turk: By the numbers

Experiments on 72,000 Amazon Mechanical Turk workers over the past year—registered through an online service called TurkPrime—produced data for many fields, a testament to the platform's growing popularity. Although Amazon may have 500,000 registered users of the online platform, studies suggest far fewer participate in research.

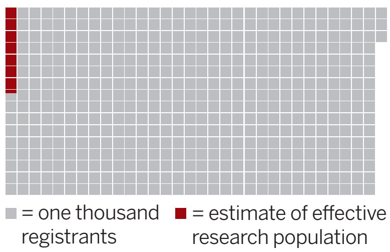
Fields of research



Published papers using Mechanical Turk



How many Turkers do research?



Corrected 9 June 2016; see full text.

Stacie Robinson, a biologist with the National Oceanic and Atmospheric Administration in Honolulu, vaccinates a Hawaiian monk seal basking on the island of Oahu.

CONSERVATION BIOLOGY

A race to vaccinate rare seals

In a first, researchers immunize wild marine mammals

By **David Malakoff**, *on Oahu, in Hawaii*

It was a rude awakening. The sunbather was snoozing on a surf-splashed ledge here on this tropical isle when a woman crept up and jabbed a hefty needle into his hip. Furious, the 2-meter-long Hawaiian monk seal snarled and lunged at his antagonist. Then, he settled back to his nap.

"All right. ... That's one more vaccinated!" exulted Tracy Mercer, a biologist with the National Oceanic and Atmospheric Administration's (NOAA's) Pacific Islands Fisheries Science Center in Honolulu.

The ambush, which occurred earlier this year on Kaena Point, Oahu's rugged western tip, is part of an unusual conservation campaign: For the first time, biologists are attempting to vaccinate a wild population of seagoing mammals in order to protect the animals from a potentially devastating virus.

It's a daunting task. Although the Hawaiian monk seal (*Neomonachus schauinslandi*) is one of the world's most endangered marine mammals, there are still some 1300 individuals scattered along the 2500-kilometer-long Hawaiian chain. For the vaccine to work, biologists must track down and give each animal two shots, weeks apart. But after years of studying ways to prevent an outbreak of phocine distemper virus (PDV), a major seal killer, NOAA researchers are optimistic that they can make wildlife health history. "It's a fascinating test case," says conservation ecologist Andrew Dobson of Princeton University, who is not involved in the effort. "People are very interested in how it is going to work."

In recent decades, PDV and other viruses in the *Morbillivirus* genus—which includes the human measles virus—have caused die-offs of tens of thousands of seals and porpoises in the Atlantic Ocean. The viruses also are circulating in the Pacific, and have been carried to Hawaii by whales and stray seals. Even Hawaii's dogs carry a potentially problematic *Morbillivirus*. So far, Hawaii's monk seals have been spared, but biologists fear an outbreak could cripple conservation efforts.

The seals are vulnerable. Geographically isolated for millions of years, they are genetically very similar and have never been exposed to PDV, studies suggest. So the seals are likely "naïve," lacking immunity to PDV and related strains, and an infection could sweep through the homogenous population.

To come up with a plan to prevent or contain an outbreak, researchers had to resolve some "major sets of unknowns" about how it might unfold, says Albert Harting, an ecologist with NOAA in Honolulu. For example, PDV typically spreads through respiratory droplets and bodily contact, so they needed to calculate how often healthy seals might encounter sick ones. They also needed to know how quickly a vaccine would take effect and how many seals they would need to immunize to prevent the disease from spreading. By tapping data from past *Morbillivirus* outbreaks in other marine mammals, vaccine tests on captive seals, and surveys of monk seal movements and behavior, the researchers developed a range of scenarios.

The models, together with a real-world vaccination drill here this past summer, indicated that it might be difficult to rapidly locate and vaccinate seals during an outbreak. Emergency vaccination might also be futile, in part because the *Morbillivirus* vaccine—developed for ferrets—takes a month or more to provide protection. But the models suggested that preventative vaccination could work. The tactic has been used infrequently in wild animals, mostly to prevent raccoons and foxes from spreading rabies, but never in free-living marine mammals.

Undeterred, NOAA last February started vaccinating seals that haul out on Oahu, which serves as a nexus between seal populations living to the northeast and southwest. "We realized that we could start on Oahu and build a firewall that might stop [an epidemic] from spreading to all three populations," says Jason Baker, a NOAA biologist in Honolulu.

Building that firewall will require vaccinating 26 of Oahu's roughly 43 known monk seals—enough to achieve "herd immunity," Baker says, and prevent essentially all the local outbreak scenarios the researchers envisioned. So far, they have vaccinated 11 Oahu seals, enough to prevent about 80% of the modeled outbreaks, as well as three seals on the neighboring island of Kauai. "That's the beauty of having such small populations," says Stacie Robinson, a NOAA ecologist in Honolulu. "You can build to big impacts pretty quickly."

Vaccination is arduous work. To reach resting seals, Robinson and her colleagues often creep across sharp lava and coral ledges, armed with a spearlike device tipped with a needle. They need to find every animal twice, once for a first dose and 4 to 6 weeks later for a booster shot. Luckily, most monk seals have numbered tags or unique markings, and volunteers help biologists track the animals.

Researchers have enough vaccine on hand to immunize 58 seals this year. Eventually, they hope to vaccinate seals on Hawaii's more remote islands, including major pupping grounds where young seals can be vaccinated relatively easily before the animals head out to sea. But that plan assumes the vaccine—which has gone off the market at times—remains available. And maintaining immunity will mean continuing to vaccinate far into the future, notes Steven Osofsky, a veterinarian with the Wildlife Conservation Society in Washington, D.C. "Essentially," he says, "these very small wild populations have to become permanently managed populations." ■

UNDERGRADUATE RESEARCH

Genuine research keeps students in science

University of Texas program boosts STEM retention and graduation rates, study finds

By **Jeffrey Mervis**

Giving college freshmen the opportunity to do research as part of their coursework significantly increases their chances of completing college and graduating with a science degree, according to a new study of a novel program at the University of Texas (UT), Austin. It's the first conclusive evidence that so-called active learning courses, which science educators have promoted for decades as a better way to teach than lectures and cookbook labs, can lower the high attrition rates in STEM (science, technology, engineering, and mathematics) fields at U.S. universities.

The relatively large size of the study, which appears in the current issue of *CBE-Life Sciences Education*, also strengthens its policy implications. President Barack Obama has challenged the country to produce 1 million more STEM-trained workers by 2020. Scaling up the UT approach nationally would be a cost-effective way to help achieve that goal, the authors argue.

Others agree. "It's a very powerful example of what can be done at our large public universities, which train a high proportion of our undergraduates," says Jo Handelsman, associate director for science at the White House Office of Science and Technology Policy in Washington, D.C., who in 2012 co-chaired an influential report on improving undergraduate education before joining the administration.

Boring introductory courses are one big reason why fewer than 40% of students who enter college as a STEM major actually earn a STEM degree within 6 years, science educators say. It's easy to imagine a better alternative: Most scientists begin their careers as an apprentice in a senior scientist's laboratory. But evidence is lacking. "Everybody knows the apprentice model works, but nobody has collected data on it," says Sarah Elgin, a biologist at Washington University in St. Louis in Missouri, who chaired a 2015 U.S. National Academies of Sciences, Engineering, and Medicine's convocation on discovery-based research.

Hard evidence has now come from UT's Freshman Research Initiative (FRI). Launched in 2005 to improve retention rates across the 50,000-student campus and, specifically, within its College of Natural Sciences, the initiative is a three-course sequence. The first class is on research methods, followed by two semesters of research in one of more than two dozen fields. As many as 40 undergraduates are introduced to the world of hypothesis-driven research by a research educator, a non–tenure track faculty member, or a postdoctoral student.

"The students have to problem-solve," says Erin Dolan, a co-author of the study and executive director of the Texas Institute for Discovery Education in Science, which oversees the FRI. "Unlike in a traditional lab course, where the next week they get a new problem, here the research stops if they don't make progress. And they are graded on attempts to solve a problem rather than results, because in science you can't anticipate the results." Tens of thousands of students have gone through the FRI, with about half of them majoring in the life sciences.

University of Texas, Austin, research educator Moriah Sandy (in gray T-shirt) works with students in a do-it-yourself diagnostics laboratory.

The program provided a large sample for the UT team, which matched 2500 students who completed the FRI courses with a group of non-FRI freshmen who had comparable demographics, abilities, and interests. That rigor allowed the authors to avoid a limitation of previous studies, which focused on students, typically upperclassmen, who applied for research-based internships or were otherwise highly motivated to pursue research experiences. The program's reach across many such disciplines makes the findings more generalizable than previous studies based on one field, or even a single course. The study also offers a more significant and objective endpoint—graduation—than many studies that ask students to assess their experience with active learning immediately after the course.

The Texas study found that 94% of FRI students graduated with a STEM degree, versus 71% of non-FRI students. In addition, 83% of FRI students graduated within 6 years, compared with only 66% of non-FRI students. There was no statistical difference in their grade point averages. Elgin applauded the authors' ability to find a causal relationship given how many factors can influence a student's pursuit of a degree and a major.

Although traditional apprenticeships clearly produce world-class science, they can be elitist, excluding many students who can't find mentors. In practice, the approach also has undervalued diversity, as women and many racial minorities can testify.

The Texas initiative offers good news on that front as well. Students from underrepresented minority groups who completed the FRI also exhibited the high retention and overall graduation rates, Dolan notes. "Because underrepresented minorities leave STEM at a higher rate, it's huge to have the same effect," she says.

UT deserves high marks for finding a scalable alternative to individual research internships, says Marcia Linn of the University of California, Berkeley, an expert on STEM education assessment. "What's really great," she says, "is that a course-based program has been designed to be so effective, to emulate the best features of an apprenticeship program with a lot more control and support."

Programs like the FRI can also be good for the bottom line, the paper notes. Not only is their cost per student much lower than the typical summer internship, but their ability to boost graduation rates means more tuition dollars for colleges. And given the demand for technically savvy workers, more STEM-trained graduates are also likely to be a boon to the U.S. economy. ■

PHOTO: THE UNIVERSITY OF TEXAS, AUSTIN

RED SQUIRRELS RISING

Scientists are striving to save a beloved native species from its disease-bearing North American cousin

By **Erik Stokstad**, *in Bangor, U.K.*

In a mossy woodland on the northwest coast of Wales, Craig Shuttleworth pulls off a dirt road and parks his battered Land Rover. Leaping over a stone wall, the tall, wiry biologist checks a trap, where a gray squirrel paces anxiously. Shuttleworth kneels, calmly slides a sturdy plastic sack around the trap's door, and blows into the cage. The squirrel, fearing the human scent, darts into the bag. The biologist quickly rolls up the sack to immobilize the animal. "I don't like doing this," he says, picking up a heavy stick worn smooth with use. "But they don't belong here."

THWACK! THWACK! The bludgeoning fractures the squirrel's head. It is another casualty in a long war against one of the world's most invasive animals, the Eastern gray squirrel. In the 140 years since the species was introduced from North America, the gray squirrel has spread across most of the United Kingdom. Along the way, it has muscled out the native red squirrel, which is considered endangered in the country.

Shuttleworth, a conservation biologist with the Red Squirrels Trust Wales, and other scientists appear to be finally turning the tide. In 2015, the trust declared the Isle of Anglesey—separated from mainland Wales by a narrow strait—free of grays, thanks to an eradication project that the 45-year-old Shuttleworth led there for 18 grueling years. This summer, culling will begin in earnest here on the mainland. "Red squirrel conservation is blossoming, because we've got proof that we can eradicate gray squirrels from the landscape," he says.

The red squirrel's range spans northern Europe to Asia, but it is especially beloved in the United Kingdom. Prince Charles, for one, thinks it should be a national mascot. Perhaps its popularity is due to Beatrix Potter, who wrote a children's book in 1903 called *The Tale of Squirrel Nutkin*, set in the Lake District. Or maybe it is memories of Tufty Fluffytail, a cartoon squirrel that for decades taught road safety to children. Whatever the reason, Brits are enamored with the creature. "People have a real pride and passion for them," says Zoe Davies, an ecologist at the University of Kent. "There's a huge amount of excitement and determination to protect the red squirrel."

In the United Kingdom, the species needs all the help it can get. Not only do gray squirrels normally outcompete the reds for food and habitat, they also carry a deadly virus called squirrelpox. Gray squirrels are immune, but when the reds catch it, they quickly succumb to the gruesome disease. There are no reliable estimates of overall populations, but grays likely outnumber reds 200 to one. Perhaps 135,000 reds live in Scotland and northern England, a fraction of earlier numbers. Farther south, a few thousand persist mainly on islands free of gray squirrels, such as Anglesey and the Isle of Wight. Conservationists have defended the northern refuges with large culls, despite adamant opposition from animal rights groups.

Yet even the most ardent advocates admit that victories are fleeting; without constant counterattacks, gray squirrels advance inexorably. Some advocates hope that the recovery of the pine marten, a relative of weasels and badgers that preys on gray squirrels, might provide long-term relief for the reds. Scientists caution, however, that much about the pine marten's resurgence and ecological impact remains unknown.

The plight of the United Kingdom's red squirrels is a cautionary tale for the rest of Europe. The gray squirrel has colonized nearly 2000 square kilometers of northwest Italy. Delayed by lawsuits from animal rights groups, biologists there missed the chance to eradicate it, giving grays an opening to spread into France and Switzerland, and ultimately to devastate red squirrels across much of their range. "The real lesson is that it's very hard to stop this invasive species," says Colin Lawton, a mammal ecologist at the National University of Ireland, Galway. "The opportunity is to catch them early before they become established."

GRAY SQUIRRELS first gained a foothold in the United Kingdom in 1876, when a wealthy silk manufacturer released a pair on his estate in Cheshire. Bigger, bolder, and easier to spot than the secretive red squirrels, the grays charmed aristocratic collectors. The most ardent enthusiast by far was the 11th Duke of Bedford, Herbrand Russell. In 1890, he set 10 free on his estate about 65 kilometers northeast of London. He also dispersed the species by giving away offspring, including six pairs as a wedding present to a friend who released them from his castle in Ireland. (All Ireland's grays are descended from those squirrels, genetic studies have shown.)

By the early 20th century, biologists knew that populations of gray squirrels were booming. And they soon noted prob-

Red squirrels, endangered in the United Kingdom, raise their distinctive ear tufts when they are alarmed.

PHOTO: © KIRK NORBURY/ALAMY STOCK PHOTO

lems: The grays were damaging young trees by stripping bark with their claws, digging up flower gardens, and raiding bird nests. "I know of more than one patriotic Englishman who has been embittered against the whole American nation on account of the presence of their squirrels in his garden," an ecologist wrote in 1931. In 1937, the U.K. Parliament banned the introduction and possession of gray squirrels.

Even earlier, scientists sounded the alarm over a troubling phenomenon: Where gray squirrels established colonies, red squirrels sooner or later vanished. Although rarely aggressive toward red squirrels and no more prolific as breeders, grays appear better adapted to broadleaf woodlands. That's primarily because grays can digest acorns, an ability they evolved in the oak-hickory forests of eastern North America. But in 1930, a University of Oxford ecologist proposed another reason for the reds' decline: The grays might be transmitting a disease.

That hunch was right. In 1981, researchers identified the culprit as a *Parapoxvirus* (the taxonomy is not settled), and experiments 20 years later confirmed that the virus kills red squirrels while sparing grays. Grays can shed the virus in scat and from scent glands, and reds somehow pick it up, perhaps through their own scent glands as they mark territory. Fleas can also spread the virus, which may happen when grays investigate red squirrel nests. Once the virus slips into a population of reds, it spreads quickly.

Gray squirrels presumably evolved immunity in North America. But red squirrels are defenseless. The virus causes weeping sores, particularly around the digits and face. The eyelids can crust over completely with scabs. Most squirrels die within a few weeks, baffling researchers. "No one really understands why it's causing mortality," says Colin McInnes, a virologist at the Moredun Research Institute in Penicuik. One idea is that sick squirrels can't eat or drink, but some dead animals have been found hydrated and nourished. Another theory behind the population collapse is that lethargic, sensory-deprived animals may be an easy target for foxes, raptors, and other predators.

Whatever the reason, the virus was decimating the red squirrel, says Peter Lurz, an independent biologist based in Randersacker, Germany, who has studied red and gray squirrels in the United Kingdom for more than 25 years. As reds succumb, gray squirrels quickly take over the habitat. When the disease is present, their range can expand by as much as 34 square kilometers per year—25 times faster than when the red squirrels are healthy, Lurz and colleagues have found.

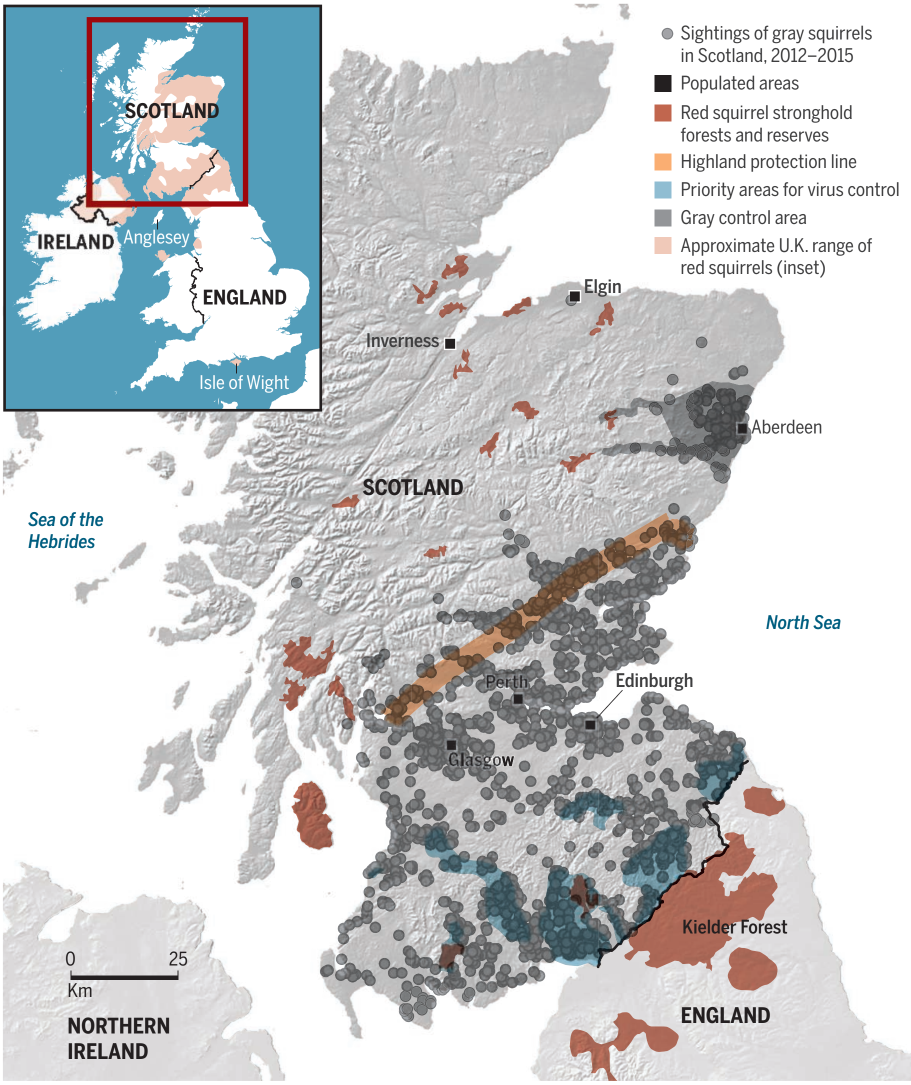
Squirrelpox's grisly symptoms boosted public sympathy for the red. "You see the animal you cherish die a horrible death," Lurz says. But the sole practical remedy—killing gray squirrels en masse—disturbs animal rights advocates. Some challenge the premise that reds, as a native species, deserve more protection than grays. Animal Aid, a U.K. animal rights organization, adds that humans have themselves to blame for worsening the red squirrels' plight; they were once considered a pest, and foresters killed untold numbers in the 1900s.

But the grays are now the real enemies. In the 1950s, a government bounty hardly dented the population. More recent eradication attempts, such as a 3-year experiment in Thetford Forest in Suffolk, also failed to push back the grays. It's not for want of trying. In Northumberland, Rupert Mitford, the 6th Baron Redesdale, has claimed to have had more than 23,000 gray squirrels killed on his estate and beyond. Prince Charles has ongoing culls on his properties in Scotland and in Cornwall, where he hopes to reintroduce red squirrels. To have a chance at success requires more than persistence. "You need a situation that is defensible," says Chris Thomas, an ecologist at the University of York. "If you can't do control to the point of exclusion, you might be throwing good money after bad."

The only unalloyed victory against the grays has been on Anglesey. The 714-square-kilometer island is fairly secure, because squirrels can reach it only by scampering

Holding the line

Most of the United Kingdom's endangered red squirrels are in Scotland. Scientists and conservationists are trying to keep invasive gray squirrels—rampant across England and Wales—out of the highlands. Grays carry squirrelpox virus, which is deadly to reds.



CREDITS: (MAP) A. CUADRA/SCIENCE; (DATA) © SAVING SCOTLAND'S RED SQUIRRELS (2016)/© SCOTTISH WILDLIFE TRUST (2016)/NORTHERN IRELAND: NATIONAL BIODIVERSITY NETWORK

across bridges. Grays first invaded it in the late 1960s. By 1998, just 40 or so red squirrels remained. Then, an avid 87-year-old conservationist named Esmé Kirby began a campaign to remove the grays and hired Shuttleworth, not long out of graduate school. By 2010, Shuttleworth's team had trapped and killed more than 6400 grays. As the population thinned, virus prevalence dropped, Shuttleworth and colleagues reported in 2014 in *Biological Invasions*. The team caught about a dozen gray squirrels in 2012 and just one the next summer. "It's amazing what they've done," Lawton says. Reds have bounced back, aided by translocation from zoos, and now number at least 700.

The next step is to defend Anglesey with a 165-square-kilometer gray-free zone on the mainland. Funding will come from Red Squirrels United, an umbrella group of 32 organizations that has several million pounds in grants from the European Union and the U.K. Heritage Lottery Fund. Anglesey is not their only point of attack. The group will train 1250 volunteers to trap and kill grays, including in northern England's Kielder Forest (see map, p. 1270), which has many reds. In addition, the group aims to secure 128 square kilometers of red squirrel habitat in Northern Ireland.

By far the largest redoubts are in Scotland and northern England, which together hold the vast majority of the population. These regions are dominated by lodgepole pine, Sitka spruce, and Scots pine—trees more to the liking of red squirrels than gray. The red squirrels may also benefit from new plantations that connect once-isolated forest patches, boosting the squirrels' genetic diversity (*Science*, 21 September 2001, p. 2246). Another major advantage for Scotland is that the squirrelpox virus didn't arrive there until 2005, so reds have been largely spared the devastating crashes seen to the south.

A collaboration called Saving Scotland's Red Squirrels (SSRS) in Edinburgh has a three-pronged defense. The Scottish Wildlife Trust and environment agencies are killing infected gray squirrels in southern Scotland to curtail viral outbreaks. Second, they are working with landowners to wipe out grays in a highland line (see map, p. 1270) in order to defend red-only habitat to the north. And staff and volunteer trappers are blitzing the region around Aberdeen, which has the only gray squirrel population north of the highland line; grays were released there in the 1970s and have not yet spread widely, which means that eradication is feasible. "Things are going well, but I don't underestimate the challenges," says Mel Tonkin, project manager of SSRS.

An unexpected piece of good news has come with the return of the pine marten. The enemy of farmers and gamekeepers because of its taste for chickens and pheasants, this cat-sized predator was nearly exterminated in the 20th century. After receiving full legal protection in 1988, the species began to rebound, and several thousand pine martens now roam the highlands. In 2007, Scottish foresters near Perth noticed that gray squirrels were less common in places where pine martens turned up. Farmers in the Irish midlands noticed a similar trend. Following up this clue, Lawton and his former Ph.D. student Emma Sheehy detailed the first known population crash of invasive gray squirrels in *Biodiversity and Conservation* in March 2014.

Red squirrels appear to be quickly recovering in the places where grays are gone, Lawton says. "It gives me faith that the red squirrel has a future." One explanation for the pattern: Gray squirrels might be easier for martens to catch. They tend to hunt on the ground, where grays search for beech nuts and acorns. Reds tend to stay in trees and nibble cones.

To boost the number of predators, the Vincent Wildlife Trust has set 20 martens free in Wales. The animals appear to be thriving; last month, researchers spotted five kits. The trust plans to release another 20 adults this fall. It is unclear how many pine martens would be needed to permanently keep gray squirrels in check, or whether they might end up ravaging reds as gray squirrels dwindle. And red squirrel advocates worry that the pine marten could be a false hope, promising a free and uncontroversial solution that could threaten funds for culling. Lawton agrees: "The real concern is that everyone stands back and assumes everything is fine."

For the time being, gray squirrel control remains in human hands. For Shuttleworth, that means more long hours prowling the woodlands in search of invaders. Walking along a dirt lane with a bloody sack slung over his shoulder, he reaches his Land Rover and tosses several furry carcasses onto a heap of traps in the back. He's had a productive day, but knows that the dozens of other traps he has set will soon claim more victims. "It's like fighting the undead," he says. "They just keep coming."

From this vantage on the mainland, Anglesey's medieval castle, built by Edward I to conquer the Welsh, can be seen across the Menai Strait. Red squirrels now have the run of the woods near the ruins. "I like the idea that my children will have a chance to see these creatures," Shuttleworth says. "We won't give up on them." The odds are daunting, but he is committed to slaying the gray invaders and safeguarding the island sanctuary. ■

Pine martens were nearly extirpated from the British Isles. Their resurgence could help control gray squirrels.

PHOTO: © PAUL GADD/ALAMY STOCK PHOTO

Published by AAAS

ENVIRONMENT

Environmental governance for all

Involving local and indigenous populations is key to effective environmental governance

By **Eduardo S. Brondizio**[1,3] *and* **Francois-Michel Le Tourneau**[2,3]

In a world increasingly thought of as overpopulated, sparsely populated spaces remain a dominant feature: ~57% of Asia, ~81% of North America, and ~94% of Australia have population densities below 1 person per square kilometer, equivalent to the population density of most of the Sahara desert (*1*). These vast, sparsely populated landscapes include rural settlements, towns, agricultural spaces, extractive economies, indigenous lands, and conservation areas. They are crucial for climate change adaptation and mitigation, from carbon sequestration to provisioning of water, food, and energy to cities. Yet governmental and nongovernmental initiatives tend to mostly pay lip service to the diverse views and needs of their populations. Without more inclusive governance, attempts to mitigate and adapt to climate change and conserve ecosystems will be compromised.

To be politically legitimate and long-lasting, incentives and regulations for better conservation and climate change mitigation must engage with the claims, rights, and knowledge of local and indigenous populations (*2*), which may be spread over immense and distant territories. The importance of local and indigenous populations in governing ecosystems and biodiversity and in meeting global climate change mitigation goals has been firmly asserted in international conventions such as the Convention on Biological Diversity [CBD article 8(j)], the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (*3*), and in agreements and commitments made at the COP21 climate meeting in Paris in 2015. How-

[1]*Department of Anthropology, Center for the Analysis of Social-Ecological Landscapes, and Ostrom Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington, IN 47405, USA.* [2]*Center for Research and Documentation on the Americas, Paris 3 Sorbonne Nouvelle University and CNRS, 75007 Paris, France.* [3]*Institut des Ameriques, 92170 Vanves, France. Email: ebrondiz@indiana.edu*

PHOTO: KEVIN FRAYER/STRINGER/GETTY IMAGES

Inclusive governance. Local and indigenous populations in many sparsely populated regions around the world—such as the Tibetan plateau—must be involved in environmental governance.

ever, efforts to overcome asymmetries in political voice, social conditions and needs, and to ensure coparticipation in decision-making have largely remained rhetorical.

Sparsely populated areas are not necessarily pristine (*4*). They are also important sectors of agriculture and animal husbandry and of various small and large extractive industries. They are home to indigenous groups, rural communities, farmers and ranchers, and small and medium-sized towns with widely varying values, priorities, and cultures. This variability must be addressed for large-scale environmental policy to advance. Although these populations have become players in the territorial governance of a wide array of ecosystem services provision, including water, food, conservation, and carbon compensation schemes, unequal benefit sharing remains the reality for many (*5*, *6*).

Sparsely populated areas often receive limited investments in human capital and infrastructure. Locally produced and extracted resources are usually transformed in distant metropolises, where value aggregation takes place, leaving behind social costs and often insolvent public administrations. Having limited access to social services and infrastructure and lacking employment opportunities, young generations migrate and circulate in search of opportunities in expanding regional urban centers. Yet sparsely populated areas are increasingly targeted to meet national and global conservation and climate mitigation goals (*7*, *8*), and local and indigenous populations, many of which are poor (*9*), are expected to take on growing responsibilities as environmental stewards.

For instance, to meet the global conservation goals specified in the CBD's Aichi Target 11, an increase of ~3% in total terrestrial and inland water protected areas is called for during the next 4 years, representing more than 3 million km^2 (*8*). Most of these areas will be in regions of Africa, Latin America, and Asia that local and indigenous populations depend on for resources, agriculture, and husbandry. This expansion implies strict restrictions on, or abandonment of, land-use systems that have in many cases coevolved and contributed to the long-term health and diversity of regional ecosystems.

Examples of questionable environmental and social outcomes of top-down conservation expansion policies abound. For example, since the 1990s, the implementation of grassland management policies imposed on 1.5 million km^2 of the Tibetan plateau by the Chinese government—including large-scale conservation areas, restriction on nomadic lifestyle and resettlements, fencing grasslands, and limiting herd size—has threatened the livelihood of millions of pastoralists whose grazing systems have coevolved with grassland species. Compounded by climate change, infrastructure development, and pollution, this type of centralized "one size fits all" approach to environmental management also leads to mixed and concerning environmental outcomes, including the possibility of speeding up the release of grassland carbon stocks and potentially threatening the water supply of Asia's largest rivers, upon which 1.6 billion people depend (*10*).

As demands for both commodities and conservation grow, so do mismatches in land use, property regimes, and governance arrangements, undermining the sustainable governance of landscapes (*11*, *12*). For instance, although indigenous, sustainable-use, and conservation areas have expanded to cover more than 40% of the Brazilian Amazon today, they are increasingly surrounded and undermined by large-scale agriculture and ranching, logging, and energy and mining extraction. Although effective in buffering deforestation, these areas are becoming islands of cultural and biological diversity, undermining their social and environmental effectiveness (*13*). New approaches are needed to reconcile conservation goals, expanding resource economies, and the role of local and indigenous populations in landscape governance (*8*, *11*, *12*).

To be effective, environmental governance solutions must involve local and indigenous populations (*2*), and national goals and international commitments must be reconciled with local and indigenous needs and cultural perspectives, as varied as they may be. Approaches and programs that bridge diverse constituencies in resource governance are emerging in many parts of the world, including rural regional governance in the United States, multifunctional landscapes in different parts of Europe, cities protecting common watersheds in the Andes, and comanagement systems in the tropics.

For instance, collaborative efforts between scientists, policy-makers, and locals are pointing to new ways of conceiving and implementing programs that combine poverty alleviation and conservation across large pastoral ecosystems of East Africa (*14*). Large international networks, such as the Global Landscape Forum led by the Center for International Forestry Research, are also bringing together a wide range of stakeholders to share ideas, propose solutions, and make commitments for the inclusive management of landscapes. Such efforts, if connected with national policies and international mitigation programs, can help to meet climate change mitigation and conservation objectives and move us closer to meeting the challenges of the United Nations' Sustainable Development Goals. ■

REFERENCES AND NOTES

1. LandScan data platform, Oak Ridge National Laboratory; web.ornl.gov/sci/landscan/landscan_data_avail.shtml.
2. J. Mistry, A. Berardi, *Science* **352**, 1274 (2016).
3. S. Díaz *et al.*, *Curr. Opin. Environ. Sustain.* **14**, 1 (2015).
4. E. C. Ellis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7978 (2013).
5. T. Sikor, Ed., *The Justices and Injustices of Ecosystem Services* (Routledge, London, 2013).
6. L. Calvet-Mir *et al.*, *Curr. Opin. Environ. Sustain.* **14**, 150 (2015).
7. J. Fairhead *et al.*, *J. Peasant Studies* **39**, 237 (2012).
8. S. H. M. Butchart *et al.*, *Conserv. Lett.* **8**, 329 (2015).
9. D. Brockington, D. Wilkie, *Philos. Trans. R. Soc. B* **370**, 20140271 (2015).
10. J. Qiu, *Nature* **529**, 142 (2016).
11. F. M. Pouzols *et al.*, *Nature* **516**, 383 (2014).
12. A. K. Duraiappah *et al.*, *Curr. Opin. Environ. Sustain.* **7**, 94 (2013).
13. F.-M. Le Tourneau, *Curr. Opin. Environ. Sustain.* **14**, 213 (2015).
14. R. Reid *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4579 (2016).

ACKNOWLEDGMENTS

F.-M.L.T. thanks the Ostrom Workshop visiting scholar program and the CNRS international mobility program. E.S.B. thanks the RIHN (Kyoto) invited scholar program.

10.1126/science.aaf5122

ENVIRONMENT

Bridging indigenous and scientific knowledge

Local ecological knowledge must be placed at the center of environmental governance

By **Jayalaxshmi Mistry**[1] and **Andrea Berardi**[2]

Indigenous land use practices have a fundamental role to play in controlling deforestation and reducing carbon dioxide emissions. Satellite imagery suggests that indigenous lands contribute substantially to maintaining carbon stocks and enhancing biodiversity relative to adjoining territory (*1*). Many of these sustainable land use practices are born, developed, and successfully implemented by the community without major influence from external stakeholders (*2*). A prerequisite for such community-owned solutions is indigenous knowledge, which is local and context-specific, transmitted orally or through imitation and demonstration, adaptive to changing environments, collectivized through a shared social memory, and situated within numerous interlinked facets of people's lives (*3*). Such local ecological knowledge is increasingly important given the growing global challenges of ecosystem degradation and climate change (*4*).

[1]Department of Geography, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK. [2]Open University, Milton Keynes MK7 6AA, UK. Email: j.mistry@rhul.ac.uk

The insights that can be gained from local indigenous knowledge are illustrated by a recent study by Klein *et al.* (*5*). The authors show that local knowledge of climate and ecological change supports the hypothesis of delayed summers on the Tibetan Plateau. This question has been vigorously debated as a result of contrasting scientific data. Interviews with Tibetan pastoralists herding livestock on a daily basis and at higher elevations found noticeable changes in seasonality, higher snowlines, and long-term changes in animal numbers, which suggested a regional warming trend underlying delayed phenological trends. This was supported by pastoralists' perceived delays in the start of summer over multidecade time scales, thereby refuting the shorter-term trends revealed by normalized difference vegetation index (NDVI) measurements and reinforcing long-term remote sensing records.

Studies with the Inuit of the Arctic region (see the photo) also show that local ecological knowledge can reveal unexpected outcomes (*6*). For example, Idrobo and Berkes have shown that the Pangnirtung Inuit of southern Baffin Island use experiential information, reflections, variations in knowledge, and sense-making to generate new understandings about the Greenland shark and its role in the Arctic marine environment (*7*). This includes knowledge about shark occurrence, habitat, and feeding behavior that is more detailed than the current scientific understanding of shark ecology. These studies show that when indigenous people seek to adapt to novel challenges such as climate change, they do not seek solutions aimed at adapting to climate change alone, but instead look for holistic solutions to increase their resilience to a wide range of shocks and stresses from various sources, some of which may have similar, or greater, negative consequences for their communities.

A growing body of published literature discusses the importance of indigenous knowledge and differing worldviews in ecosystem science and management (*8*, *9*). Yet there is still a tendency among the scientific community to assimilate local ecological knowledge within Western worldviews of managing nature. Examples include community monitoring, reporting, and verification as part of the REDD+ policy (*10*) and the use of indigenous

Respecting local knowledge. To be successful, efforts to protect biodiversity and respond to climate change must engage with local communities and start from the perspective of indigenous knowledge of local ecology. In this photo, Inuit hunters share the fresh meat of a newly shot seal, Baffin Island, Nunavut, Canada.

Published by AAAS

fire practices for carbon abatement. Both of these are attempting to institutionalize indigenous knowledge into existing environmental governance structures that are dominated by an incentive- and market-based approach to climate change mitigation (*4*). In the case of fire management, the accounting and metrics involved in monitoring new emissions-reducing programs is a dramatic shift from how local knowledge is usually embedded in practice, place, and dynamic decision-making (*11*). This approach risks further marginalizing indigenous people.

A major reason for the limited engagement with indigenous knowledge is the persistence of epistemological differences, and the associated politics of representation, within the social and governance context. Local ecological knowledge is seen as subjective, arbitrary, and based on qualitative observations of phenomena and change. Scientific knowledge, by contrast, is viewed as objective and rigorous, with precise measuring and empirical testing of events and trends confirming credibility and legitimacy. Attempts to evaluate local ecological knowledge thus often use scientific methods to prove its validity. However, all forms of knowledge, including scientific knowledge, are produced by socially situated actors and are value-laden (*12*).

"Indigenous knowledge systems, and the processes for their evolution over time, can support rapid adaptation to complex and urgent crises."

Furthermore, the scientific approach, with its imperative for precise categorization and abstract generalization, rapidly loses its ability to provide useful guidance to the general public when faced with increasingly complex situations typified by uncertainty, nonlinear dynamics, and conflicting perspectives (*13*). Indigenous knowledge can circumvent some of these problems by generating a systemic understanding of a complex environment and integrating a large number of variables qualitatively over an extended period of time. Through collective and adaptive dialogue, indigenous knowledge can lead to simple rules that can be easily remembered and locally enforced through social means (*14*).

Conservation and development ideologies worldwide are heavily influenced by politically dominant Western agendas, and the structures in which indigenous knowledge is used and applied are determined by science. The danger is that in these places, indigenous knowledge will change in its use and application, and, most critically, in its ability to deal with complexity. For example, the institutionalization of indigenous fire management has focused on protective early dry-season burning at the expense of regular and sometimes opportunistic burning throughout the dry season and in the wet season (*11*). This could lead to a loss in the complexity of fire knowledge, amplified by a general loss of traditional knowledge (especially among young people), which has serious implications for future indigenous cultures and their linked ecosystems.

Indigenous knowledge systems, and the processes for their evolution over time, can support rapid adaptation to complex and urgent crises (*15*). Rather than encouraging these knowledge systems to become more "scientific," we urge a respectful acknowledgment of their distinctiveness and epistemology (*16*). We suggest that any effort to solve real-world problems should first engage with those local communities that are most affected, beginning from the perspective of indigenous knowledge and then seeking relevant scientific knowledge—not to validate indigenous knowledge, but to expand the range of options for action. This would make scientific knowledge more acceptable and relevant to the societies that it seeks to support, while critically promoting social justice and establishing self-determination as a key principle of engagement. ■

REFERENCES AND NOTES

1. W. Walker *et al.*, *Carbon Manage.* **5**, 479 (2015).
2. J. Mistry, A. Berardi, C. Verwer, G. de Ville, *Up-Scaling Support for Community Owned Solutions. A Project COBRA Report for Policy Makers* (2015); http://projectcobra.org/up-scaling-support-for-community-owned-solutions.
3. J. Mistry, in *International Encyclopedia of Human Geography*, R. Kitchin, N. Thrift, Eds. (Elsevier, 2009), vol. 5, pp. 371–376.
4. J. D. Ford *et al.*, *Nat. Clim. Change* **6**, 349 (2016).
5. J. A. Klein *et al.*, *Glob. Environ. Change* **28**, 141 (2015).
6. P. Bates, *Arctic Anthropol.* **44** (no. 2), 87 (2007).
7. C. J. Idrobo, F. Berkes, *Hum. Ecol.* **40**, 405 (2012).
8. E. S. Brondizio, F.-M. Le Tourneau, *Science* **352**, 1272 (2016).
9. G. Mace, *Science* **345**, 1558 (2014).
10. REDD+ is a United Nations climate change mitigation program for countries to reduce emissions from deforestation and forest degradation, and foster conservation, sustainable management of forests, and enhancement of forest carbon stocks.
11. J. Mistry, B. Bilbao, A. Berardi, *Philos. Trans. R. Soc. B* 10.1098/rstb.2015.0174 (2016).
12. K. Weiss, M. Hamann, H. Marsh, *Soc. Nat. Resour.* **26**, 285 (2013).
13. F. Berkes, M. K. Berkes, *Futures* **41**, 6 (2009).
14. E. Gómez-Baggethun, E. Corbera, V. Reyes-García, *Ecol. Society* **18**, 72 (2013).
15. J. D. Ford, G. McDowell, T. Pearce, *Nat. Clim. Change* **5**, 1046 (2015).
16. N. Klenk, K. Meehan, *Environ. Sci. Policy* **54**, 160 (2015).

ACKNOWLEDGMENTS

We thank two anonymous referees for their supportive and useful comments. This Perspective was inspired by more than 20 years of engagement with indigenous communities in South America and research carried out in Project COBRA (www.projectcobra.org), funded by the Environment Programme, Management of Natural Resources, DG Research and Innovation, European Commission 7th Framework.

10.1126/science.aaf1160

IMMUNOTHERAPY

Outsourcing the immune response to cancer

Healthy donors provide T cells that target neoantigens in cancer patients

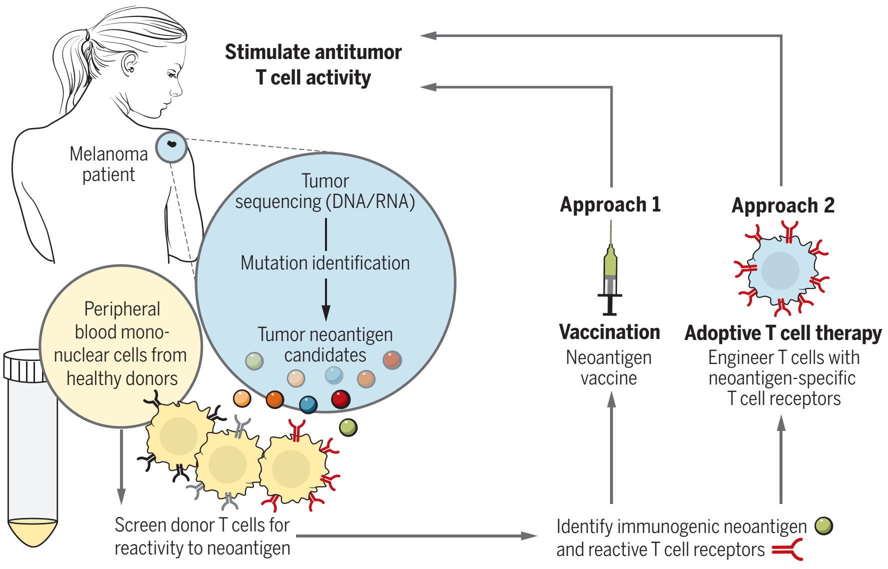
By **Mahesh Yadav** *and* **Lélia Delamarre**

One challenge in the field of cancer immunotherapy is to improve patient responses to current immune therapies that enhance the capacity of T cells to kill tumors by blocking inhibitory pathways called immune checkpoints. The recent identification of mutated proteins in tumors, called neoantigens, as the main targets of effective antitumor immunity offers the opportunity to design therapies that stimulate highly focused antitumor T cell responses (*1*). On page 1337 of this issue, Strønen *et al.* (*2*) report that a much greater frequency of tumor mutations are immunogenic than initially estimated, and that healthy donors can provide T cells that are reactive to these neoantigens. The findings point to a new individualized approach for expanding immunotherapies.

Cancers can accumulate hundreds of mutations. How many of these mutations are immunogenic is unclear. For a neoantigen to be immunogenic, the mutant protein has to be processed, and the resulting mutant peptide needs to be presented by major histocompatibility complex (MHC) molecules. Also, the MHC-mutant peptide complex must be recognized by the host's T cells. Based on preclinical evidence and immune monitoring of cancer patients, it is typically thought that only a very small fraction of expressed tumor mutations (<1%) are immunogenic in cancer patients (*3*, *4*). However, the tumor microenvironment presents major challenges to antitumor T cell responses. In addition to inhibiting existing T cell responses, the microenvironment can prevent adequate T cell priming or induce immune system tolerance to the tumor (*5*). How,

Department of Cancer Immunology, Genentech, MS 231B, 1 DNA Way, South San Francisco, CA 94080, USA. Email: delamarre.lelia@gene.com

then, can one determine the immunogenicity of tumor mutations? Strønen *et al.* asked whether naïve T cells from healthy individuals with overlapping MHC haplotypes could recognize immunogenic mutated peptides from cancer patients. The authors conducted whole exome and RNA sequencing of tumors from three melanoma patients to identify single nucleotide variants mapped to coding regions. To narrow down the number of neoantigen candidates, they next used an MHC class I peptide binding prediction algorithm as a selection criteria for immunogenicity (*6*). Of the 249 neoantigen candidates, 57 with the highest predicted MHC class I binding affinity were then tested for immunogenicity using healthy donor peripheral blood mononuclear cells. These cells were cultured with autologous dendritic cells transduced with mRNA encoding multiple (10 to 21) neoantigen candidates. They found that 11 out of 57 selected neoantigen candidates from three melanoma patients elicited reactivity in T cells from multiple healthy individuals. By contrast, T cell responses to only two of the total predicted neoantigens were detected in the tumor-infiltrating lymphocytes isolated from the cancer patients. This suggested that most immunogenic mutations go undetected when a patient's T cells are screened for tumor reactivity. Strønen *et al.* further showed that neoantigen-specific T cells from healthy donors could recognize and kill melanoma cells harboring the relevant mutation in vitro. These data indicate that the frequency of immunogenic neoantigens has been largely underestimated and that tumors with a low mutation load could have greater immunogenic potential than originally thought.



Improving immunotherapy. Use of healthy donors' T cells can improve the identification of immunogenic neoantigens for the development of neoantigen-based vaccines or T cell–based adoptive cell therapy.

The current gold standard for predicting peptide immunogenicity relies on predicting MHC class I binding affinity to candidate peptides using in silico approaches (*4*, *6*). Although this strategy substantially enriches for immunogenic neoantigens, the incidence of false positives is high (80% of the predicted positives were false in the study of Strønen *et al.*). The authors show that the neoantigens recognized by donor T cells made a stronger complex with MHC and displayed a longer half-life as compared to the peptides that were not recognized. Thus, they demonstrate that neoantigen selection could be improved by adding peptide-MHC complex stability as an immunogenicity criterion. Other factors, such as position of the mutated amino acid as a predictor of T cell receptor recognition or the nature of the amino acids in the T cell receptor-contact region of the peptide, have also been associated with immunogenicity and should be explored to further improve the accuracy of prediction algorithms (*3*, *7*).

A key question is why the T cell repertoire of healthy individuals has broader reactivity to neoantigens than does that of cancer patients. Strønen *et al.* assessed existing T cell responses in the melanoma patients by focusing on the reactivity of tumor-infiltrating lymphocytes toward the 249 candidate neoantigens. However, the authors did not examine the repertoire of naïve T cells in the circulating blood of cancer patients as they did for the healthy donors. This analysis would have helped determine if the absence of tumor-reactive T cells in patients results from the lack of in vivo T cell priming or from T cell tolerance of the tumor. Such information would enable development of an appropriate neoantigen-based therapeutic regimen depending on the outcome. Evidence of broader antitumor T cell repertoire in a cancer patient's blood in comparison to the tumor-infiltrating lymphocytes would suggest a defect in priming. Such a defect could potentially be addressed through vaccination. Indeed, vaccination of melanoma patients with neoantigen peptides increased the breadth of their immune responses (*8*). The approach of Strønen *et al.* for identifying neoantigens that elicit T cell responses could further help prioritize peptide candidates for vaccination (see the figure).

An alternative approach to targeting neoantigens is by transferring neoantigen-specific T cells directly into the host by adoptive T cell therapy. Typically, this approach encompasses isolation of tumor-infiltrating lymphocytes from a patient, ex vivo selection of tumor-specific T cells, and expansion into a large number of cells for therapeutic infusion. Several studies suggest that neoantigens are important targets for successful adoptive T cell therapy (*9*, *10*). Strønen *et al.* asked whether it is feasible to transfer neoantigen reactivity to a patient's T cells by transferring T cell receptors from the healthy donor's T cells that recognize neoantigens. The authors show that T cell receptors in donor T cells that recognize neoantigens in vitro can be sequenced, expressed in T cells, and effectively recognize (and therefore facilitate the killing of) tumor cells in vitro. Thus, this approach could provide an excellent means to systematically identify allogeneic T cell receptors directed against neoantigens and support the therapeutic exploitation of allogeneic T cell receptor–engineered T cells for adoptive cell therapy.

The rapid progress of neoantigens as ideal targets for eliciting effective anticancer immune responses has provided further hopes for boosting cancer immunotherapy. However, only a small proportion of neoantigens fully meet the criteria and a systematic approach for identifying immunogenic neoantigens in a high-throughput manner is needed to advance neoantigen-based approaches in cancer immunotherapy. The findings of Strønen *et al.* now put us more firmly on that path. ■

REFERENCES

1. T. N. Schumacher, R. D. Schreiber, *Science* **348**, 69 (2015).
2. E. Strønen *et al.*, *Science* **352**, 1337 (2016).
3. M. Yadav *et al.*, *Nature* **515**, 572 (2014).
4. N. Van Rooij *et al.*, *J. Clin. Oncol.* **31**, e439 (2013).
5. D. S. Chen, I. Mellman, *Immunity* **39**, 1 (2013).
6. M. Nielsen, M. Andreatta, *Genome Med.* **8**, 33 (2016).
7. D. Chowell *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E1754 (2015).
8. B. M. Carreno *et al.*, *Science* **348**, 803 (2015).
9. S. A. Rosenberg, N. P. Restifo, *Science* **348**, 62 (2015).
10. E. Tran *et al.*, *Science* **344**, 641 (2014).

10.1126/science.aag1547

ILLUSTRATION: K. SUTLIFF/*SCIENCE*

ORGANIC CHEMISTRY

Three catalysts for activating carbon-hydrogen bonds

A general approach merges photoredox-mediated hydrogen-atom transfer and nickel catalysis to make C–C bonds

By **Corinne Fruit**

Transition metal–catalyzed arylation of C–H bonds has been intensively studied for forming C–C bonds in complex-molecule synthesis (*1*). An acidic C–H bond (for example, one near a double bond or an O atom) is cleaved to form a carbon–metal bond, which then couples to arene. Many of these organometallic species can be generated catalytically. Much less research has dealt with unreactive nonacidic sp^3 C–H bond functionalization (*2*). On page 1304 of this issue, Shaw *et al.* (*3*) report an efficient and general method that focuses on arylation of sp^3 C–H bonds at carbon atoms adjacent to amines and to cyclic ethers by combining nickel, visible-light photoredox, and hydrogen-atom transfer (HAT) catalysis.

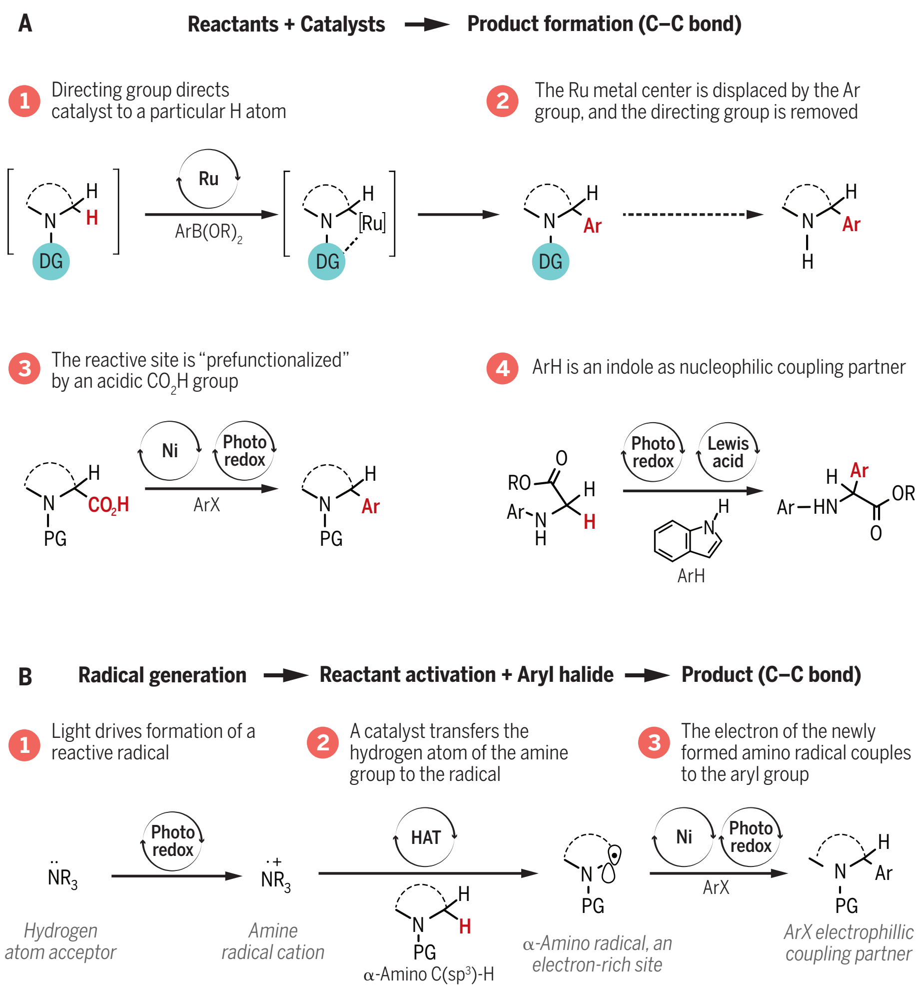
It is more difficult to achieve C–H functionalization with sp^3 C–H bonds than with sp^2 C–H bonds because the former are less acidic. Nevertheless, α-amino and α-oxy sp^3 C–H bond arylation proceeded readily because the heteroatom induces electronic asymmetries (an effect previously exploited in radical-mediated activation processes), whereas coordination to the catalyst led to strong directing effects (*4*).

Organic compounds usually have several sp^3 C–H bonds that could be potential sites for C–H functionalization, so selectivity is critical. Often-used strategies for controlling product selectivity in catalytic reactions are the installation of a directing group onto the substrates (*5*) or the use of prefunctionalized compounds (such as carboxylic acids) (*6*). Both of these strategies have been applied to amine substrates (see the figure, panel A), but selectivity comes at the price of having to remove these groups from the product.

Shaw *et al.* instead used a combination of three catalysts, which enabled them to define a broad scope of both viable substrates and electrophilic coupling partners. The latter provides more options for the introduction of diversity into target compounds and could lead to the generation of compounds with increased novelty.

Visible-light photoredox catalysis has been recognized as a powerful and environmentally friendly activation strategy in chemical transformations and catalytic chemical processes (*7*). This approach relies on the ability of a light-sensitive catalyst to engage in single-electron transfer processes with organic substrates upon photoexcitation. As a result, highly reactive radical-ion intermediates are generated with often unusual reactivities under mild reaction conditions. Amines are very good electron donors that readily undergo single-electron oxidation to yield radical cations. A combination of visible-light photoredox and Lewis acid was previously described for the α-amino C–H bond functionalization with this radical cation (*8*) but was limited to glycine amino acid (or dipeptide) and indole derivatives as the nucleophilic coupling partners (see the figure).

The striking feature of the merger of HAT and photoredox catalysis described by Shaw *et al.* is that it generates α-amino radicals (see the figure, panel B). As highly nucleophilic radicals, α-amino radicals will undergo C–H arylation with electrophilic aryl halides under nickel catalysis. By the simple addition of a suitable hydrogen-atom acceptor in the reaction (amine HAT catalyst), the synergistic effect of diverse activation modes described by Shaw *et al.*



Catalytic coupling to form C–C bonds from C–H bonds. (**A**) Previously reported catalytic strategies for arylation of amines at the C–H bond of its α-carbon. Abbreviations: DG, directing group; PG, protecting group; R, alkyl; Ar, aryl; and X, iodide or bromide. (**B**) The route of Shaw *et al*. that uses three catalysts; only the protected amine reactant is shown, but the method also works with cyclic ethers.

GRAPHIC: V. ALTOUNIAN/SCIENCE

Normandie Univ., INSA Rouen, UNIROUEN, CNRS, COBRA, 76000 Rouen, France. Email: corinne.fruit@univ-rouen.fr

notably expands the substrate's scope and decreases the amounts of the catalysts needed. It allows structural diversity to be incorporated into complex heterocycles and for late-stage functionalization of highly valuable scaffolds such as cyclic amines and tetrahydrofuran skeletons, useful in drug discovery. The reported C–C coupling reaction displays high functional group tolerance with good yields. Finally, Shaw *et al.* developed a notable straightforward synthesis of a 2,5-diarylated pyrrolidine, a widespread structural feature of natural and designed biologically active molecules (*9*), from readily available proline derivatives by using sequential C–H and decarboxylative arylation.

The powerful strategy reported by Shaw *et al.* could give rise to exciting developments in enantioselective transformation by using chiral photocatalysts (*10*) or ligands (*11*) and prochiral substrates bearing different functional groups at the α-carbon position. An asymmetric photoredox catalysis could offer new opportunities for the synthesis of nitrogen-containing chiral molecules and the generation of quaternary stereogenic centers. In addition, extension of this method to an intramolecular version

> ***"The striking feature of the merger of HAT and photoredox catalysis described by Shaw* et al. *is that it generates α-amino radicals."***

of the sp^3 C–H arylation of cyclic amines could provide access to tricyclic heterocycle compounds. This class of compounds, such as pyridoindolizidine derivatives (*12*), are of pharmaceutical interest as antagonists for nicotinic acetylcholine receptors. ■

REFERENCES

1. A. K. Pitts, F. O'Hara, R. H. Snell, M. J. Gaunt, *Angew. Chem. Int. Ed.* **54**, 5451 (2015).
2. O. Baudoin, *Chem. Soc. Rev.* **40**, 4902 (2011).
3. M. H. Shaw, V. W. Shurtleff, J. A. Terrett, J. D. Cuthbertson, D. W. C. MacMillan, *Science* **352**, 1304 (2016).
4. K. Campos, *Chem. Soc. Rev.* **36**, 1069 (2007).
5. S. J. Pastine, D. V. Gribkov, D. Sames, *J. Am. Chem. Soc.* **128**, 14220 (2006).
6. Z. Zuo *et al.*, *Science* **345**, 437 (2014).
7. C. K. Prier, D. A. Rankic, D. W. C. MacMillan, *Chem. Rev.* **113**, 5322 (2013).
8. S. Zhu, M. Rueping, *Chem. Commun.* **48**, 11960 (2012).
9. F. Bellina, R. Rossi, *Tetrahedron* **62**, 7213 (2006).
10. H. Huo *et al.*, *Nature* **515**, 100 (2014).
11. Q. M. Kainz *et al.*, *Science* **351**, 681 (2016).
12. J. L. Kristensen *et al.*, *ACS Med. Chem. Lett.* **6**, 472 (2015).

10.1126/science.aaf8923

GENOMICS

A federated ecosystem for sharing genomic, clinical data

Silos of genome data collection are being transformed into seamlessly connected, independent systems

The Global Alliance for Genomics and Health*

Early data-sharing efforts have led to improved variant interpretation and development of treatments for rare diseases and some cancer types (*1*–*3*). However, such benefits will only be available to the general population if researchers and clinicians can access and make comparisons across data from millions of individuals.

Despite much progress, genomic and clinical data are still generally collected and studied in silos: by disease, by institution, and by country. Regulatory data-privacy requirements do not seamlessly lend themselves to the secure sharing of data within and across institutions and countries (*4*). Current practices in research and medicine hinder the sharing of data in ways that tangibly recognize an individual's contributions. Tools and analytical methods are nonstandardized and incompatible, and the data are often stored in incompatible file formats (see the figure). If we stay this course, the likely outcome will be an assortment of balkanized systems akin to those developed for U.S. electronic health records, which, although designed to advance human health by sharing clinical data across institutions, have by all measures fallen short of that goal because of a lack of interoperability.

POLICY

A FEDERATED DATA ECOSYSTEM. The Global Alliance for Genomics and Health (GA4GH) was established in 2013 to enable responsible and effective sharing of genomic and clinical data in a way that is as simple as using the World Wide Web. GA4GH, which now brings together hundreds of individuals and organizations, was built on the hypothesis that the data underlying genomic medicine must be federated. That is, whereas data may be distributed across many databases and computers around the world, they must be virtually connected through software interfaces that allow seamless, authorized access. In contrast to large centralized data repositories, a federated system will allow legal data control to remain within the originating jurisdiction (see the figure). International consortia such as the International Cancer Genome Consortium (ICGC) have already adopted federated databases because the model allows local databases to maintain autonomy (*5*).

TOOL DEVELOPMENT AND USE. As a first step, the GA4GH Regulatory and Ethics Working Group (REWG) developed a framework document that provides basic principles and core elements for responsible data sharing (*6*, *7*) and is founded on Article 27 of the 1948 *Universal Declaration of Human Rights* (*8*). This focus on human rights represents a paradigm shift with respect to data sharing, as most previous discussions focused solely on protection from harm without acknowledging the right to benefit from the fruits of scientific and medical advances. In practical terms, increased data sharing will enable researchers to make better predictions about disease risk, prevention, and treatment by virtue of having access to larger data sets. And through data exchanges that link the clinical and research communities, clinicians will be able to make better precision medicine decisions for individual patients.

Additionally, the Data Working Group (DWG) has developed a standardized application programming interface (API), which offers a defined protocol to allow disparate technology services of institutions around the globe to communicate with one another to exchange genotypic and phenotypic information.

The API and the framework document are being used in demonstration projects spearheaded by GA4GH members.

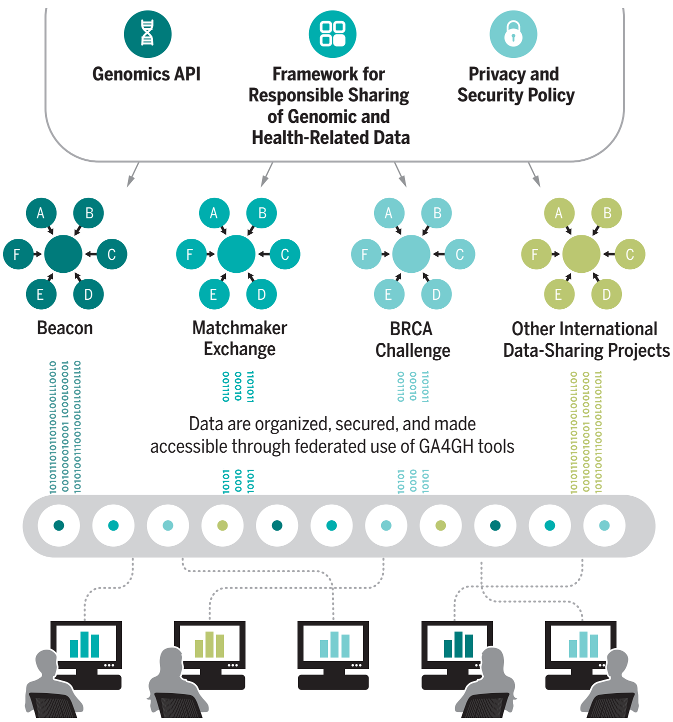
Beacon Project. The Beacon Project (http://ga4gh.org/#/beacon) is developing an open technical specification for sharing genetic variant data sets collected from large-scale population-sequencing projects, clinical diagnostic settings, and variant curation efforts available to the community. A beacon is a Web-accessible service that allows data sets to be queried for the presence or absence of a specific allele. A user of a beacon can ask it questions of the form, "Have you observed

*See supplementary materials for full author list. Corresponding author. Email: peter.goodhand@genomicsandhealth.org

this nucleotide (e.g., C) at this genomic location (e.g., position 32,936,732 on chromosome 13)?" to which the beacon must respond with either "yes" or "no." In the 2 years since the project's launch, 23 organizations have lit more than 60 beacons serving more than 200 data sets. The data sets served through beacons may be queried individually or in aggregate via the Beacon Network, a federated search engine (http://beacon-network.org/#). Currently, all Beacon users must agree to a single set of data-use conditions. However, work is under way to allow Beacon users to choose from a predetermined set of conditions that restrict potential data use on the basis of the consent of individuals represented in the data (*9*).

In contrast to traditional "all-or-nothing" approaches to sharing data [e.g., open or password-protected access to variant call format (VCF) files], Beacon uses a tiered access approach, in which increasingly detailed information is made available to users at more stringent levels of authentication and authorization and with a formal specification of data-use conditions. A registered access level that would fall between open and controlled access is under development.

A federated data ecosystem. To share genomic data globally, this approach furthers medical research without requiring compatible data sets or compromising patient identity.



By adopting a federated model, Beacon overcomes the inefficiency and expense experienced when data generators must transfer whole copies of their data sets into a single, centralized repository. The federated approach also circumvents the often-prohibitive privacy and security risks that arise when such transfers force data to cross institutional and, sometimes, national or continental boundaries. Also, because Beacon is compatible with any underlying representation of alleles or allelic annotations, it is not limited by particular file formats. Finally, Beacon allows data discovery without exposing identifiable information, because it does not require data generators to share fully described data representations or annotations.

BRCA Challenge. The BRCA Challenge aims to advance understanding of the genetic basis of breast, ovarian, and other cancers that are driven by germline variants in *BRCA1* and *BRCA2*. The project's first product is the BRCA Exchange (http://brcaexchange.org), a publicly accessible Web portal that provides a simple interface for patients, clinicians, and researchers to access curated, expert interpretations of *BRCA1/2* genetic variants, as well as supporting evidence. An expanded research arm of the portal was recently launched to allow any Web user to access data from the original submitters. A third tier of access will be made available to registered users who require access to potentially identifiable case-level data and are working on variant interpretation. In addition to developing the BRCA Exchange, the BRCA Challenge team members are working to understand the liability concerns faced by federated databases of this kind, such as misclassifications or failure to regularly validate and update classifications.

Matchmaker Exchange. Matchmaker Exchange (MME) is a collaborative effort of consortia, including members of the International Rare Diseases Research Consortium (www.irdirc.org/) and related laboratories in the rare disease space, where the majority of cases studied lack a clear etiology after initial analysis (*10*). Given a suspicious variant in a candidate disease gene, matching two cases that share the variant, as well as an overlapping phenotype, may be sufficient evidence to causally implicate the gene. To facilitate such discovery, researchers in the rare disease community have established a series of platforms that allow users to identify cases with phenotypes and disrupted genes in common. MME was established to connect rare disease databases, such that a query to one would enable searches of the others, without having to deposit data into each one.

At this time, three Matchmaker Services (GeneMatcher, Phenome Central, and DECIPHER) have implemented the MME API. To ensure accurate comparison of patients assessed by different clinicians, similar phenotypes are determined by matching identical or ontologically similar terms with the Human Phenotype Ontology. MME users must deposit their data into an existing MME service, and tools on the MME website (http://matchmakerexchange.org) help guide users toward the database that is most appropriate for a given case. Although the system is currently geared toward clinicians and researchers, the team is also working with patients to establish patient-led matchmaking endeavors with support from such organizations as Free the Data and the ClinGen Genome Connect Registry.

Matchmaking has already led to the diagnosis of several previously undiscovered rare diseases (*10*). Successful matching will increase considerably as the volume of cases connected through MME increases. Additionally, MME will soon enable "hypothesis-free" matching in which the genotype aspects of matching can occur by direct query of variants within a VCF that meet certain criteria, even if no candidate gene has been labeled as such. This will require MME services to support queries of entire genomic data sets.

Finally, with input from the REWG, MME has developed a two-tiered informed-consent policy to define the type of consent needed for using MME and when no consent is needed. If the data are associated with a unique or sensitive phenotype or with sequence-level data, consent from the patient is required to share it for research purposes. However, if only standard phenotype terms and candidate gene names are used, consent to clinical care allows for matchmaking. Still, challenges remain in balancing discovery with privacy and data protection.

A variety of issues arise when data must cross multiple communities (e.g., patient privacy, distinct international laws, individual academic success in gene discovery, user authentication and consistent standards for data exchange across distinct databases). Although GA4GH has been convening stakeholders to address these challenges, more groups and data sets must still be brought in.

REMAINING CHALLENGES. Shringarpure and Bustamante (*11*) used simulations to show that, in some scenarios, querying a public beacon for as few as 250 variants al-

GRAPHIC: G. GRULLÓN/*SCIENCE*

ready known to be present in an individual's genome could reveal information distinctive to that individual. GA4GH members have been developing solutions to this potential security breach since the project's inception, including aggregating data among multiple beacons, tracking usage to restrict systematic searches and introducing tiers of secured access that require users to be authorized for data access—but these necessarily limit the scope of information that can be shared widely. Innovative policy and regulatory measures, as well as technological solutions, are needed to securely handle individual genomic and clinical data.

A second challenge is scalability. For every problem there will be domain-specific challenges that may require uniquely applicable tools. For instance, the field of dementia research may demand new solutions that integrate data from brain MRI technology. Furthermore, it is expected that individual fields will have previously developed standards, which may demand that GA4GH adapt its existing solutions in order to be compliant. Applying existing GA4GH approaches in new contexts will require solutions that are easily portable, customizable, and interoperable. GA4GH must also focus on solutions that can benefit many different patient groups, jurisdictions, health systems, and environmental and socioeconomic realities, such as interoperability with mobile devices, which are now broadly available even in developing nations. Open technology, built-in interoperability, and ease of use of data-sharing tools are essential.

Data sharing has inherent costs related to data curation, hosting, and computation. Hoopen *et al.* described substantial costs of post-data curation, leading to a proposal for lower-cost submitter-driven annotation as a more sustainable curation solution (*12*). Many databases currently recover costs through user fees (*13*), creating either a need to charge and share revenue or a two-tiered system that may limit some downstream users from accessing the full complement of information. Member projects, such as ICGC's PanCancer Analysis of Whole Genomes (PCAWG), have implemented federated cloud-based solutions that bring the cost of analyzing a single sample from U.S. $200 by using traditional academic high-performance computing models to under U.S. $20 per sample. Cloud-based approaches also have the benefit of being compatible with some country-specific legal frameworks (*14*). Several business models to support genomics big data research have been proposed, including a subscription model, which may inherently limit access, and a "freemium" model, which charges not for data access but for associated services, such as curation and interpretation (*15*).

Notwithstanding emergence of new business models for private and public sector partnerships to support some data-sharing costs, government agencies may need to support some features of the ecosystem (e.g., curation) so that clinicians and patients have access to as much free, curated data as possible. In addition to economic incentives, more can be done to establish greater academic recognition for data sets through citations and microattribution, in which quantitative credit is attached to every data-use accession (*16*, *17*).

Finally, ensuring engagement among the entire global community is necessary from a social justice and medical perspective, although this will likely require distinct legal, cultural, and business models. In some countries, health care and research organizations are interested in GA4GH as a means to link nascent national efforts in precision medicine with other international groups, such as the Brazilian Initiative on Precision Medicine (www.fcm.unicamp.br/gtc/evento/1/trabalho/8). Training and infrastructure needs related to data storage, management, security, and policies are common to many jurisdictions. Technology and economic incentives can make it possible for an international, federated network of genomic and clinical data to become a network for learning that will illuminate causes of disease and potential interventions for prevention and treatment. ■

REFERENCES

1. B. Coste *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4667 (2013).
2. H. Davies *et al.*, *Nature* **417**, 949 (2002).
3. J. C. Cohen, E. Boerwinkle, T. H. Mosley Jr., H. H. Hobbs, *N. Engl. J. Med.* **354**, 1264 (2006).
4. The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation, Council of Canadian Academies, *Accessing Health and Health-Related Data in Canada* (CCA, Ottawa, ON, 2015).
5. T. J. Hudson; The International Cancer Genome Consortium, *Nature* **464**, 993 (2010).
6. The Global Alliance for Genomics and Health, Framework for Responsible Sharing of Genomic and Health-Related Data (GA4GH, Toronto, 2014); https://genomicsandhealth.org/framework.
7. B. M. Knoppers, *HUGO J.* **8**, 3 (2014).
8. United Nations General Assembly, *Universal Declaration of Human Rights* (United Nations, 1948).
9. S. O. M. Dyke *et al.*, *PLOS Genet.* **12**, e1005772 (2016).
10. A. A. Philippakis *et al.*, *Hum. Mutat.* **36**, 915 (2015).
11. S. S. Shringarpure, C. D. Bustamante, *Am. J. Hum. Genet.* **97**, 631 (2015).
12. P. ten Hoopen *et al.*, Database, bav126 (2016).
13. C. Ember, R. Hanisch, "Sustaining domain repositories for digital data: A white paper" (The Interuniversity Consortium for Political and Social Research, 2013).
14. L. D. Stein, B. M. Knoppers, P. Campbell, G. Getz, J. O. Korbel, *Nature* **523**, 149 (2015).
15. P. E. Bourne, J. R. Lorsch, E. D. Green, *Nature* **527**, S16 (2015).
16. Editors, *Nat. Genet.* **39**, 423 (2007).
17. Editors, *Nat. Genet.* **390**, 931 (2007).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1278/suppl/DC1

10.1126/science.aaf6162

PHOTO: AP PHOTO/PETER DEJONG

CLIMATE CHANGE

Global adaptation after Paris

Climate mitigation and adaptation cannot be uncoupled

By **Alexandre K. Magnan** *and* **Teresa Ribera**

POLICY

Besides achieving major decisions on greenhouse gas (GHG) emissions mitigation, the 2015 Paris climate change Agreement (*1*) also initiated a process to "establish a global goal on adaptation" (Article 7.1), a crucial step that encourages parties to the agreement to go beyond the restrictive and historic funding-focused lens that structured United Nations Framework Convention on Climate Change (UNFCCC) talks on adaptation until now (*2*–*4*). Suggesting that global adaptation is as important as global mitigation is an important shift in international climate negotiations that highlights the importance of not uncoupling 21st-century

Dutch coast line reinforced in anticipation of rising water. Sand supplementation uses natural currents to push sand onto coastal dunes and dikes—in Petten, Netherlands, 7 million cubic meters (247 million cubic feet) of sand were dumped in 2008 in front of a sea-dike.

mitigation and adaptation storylines. After all, one cannot define the "well below +2°C" long-term temperature goal as sustainable without providing evidence on societies' ability to adapt to the unavoidable impacts of such warming (*5*). Although this represents great progress, we discuss three key challenges around the development of a global adaptation framework within the UNFCCC: defining a global goal, identifying tracking criteria, and anticipating political barriers. A major underlying condition is that the framework must make sense from both a negotiation and a scientific perspective.

For the first time, parties are encouraged to build a collective understanding of what adaptation means, notably through definition of references and tools to capture, track, and aggregate national adaptation efforts (Art. 7.14 and Decision 1/CP.21 paragraph 43b). A more comprehensive framework for global adaptation can help answer a crucial question that parallels the one on global mitigation: Are we as humankind on track to adapt to climate change?

UNDERLYING RATIONALE. Before Paris, many international scientific efforts, such as the Intergovernmental Panel on Climate Change (IPCC) syntheses, highlighted the importance of adaptation due to the irreversibility of some climate change impacts. Almost exclusively focusing on local-to-national approaches, they raise three main conclusions. First, although some territories are at the frontline of climate change impacts and will be affected first (e.g., small islands, arctic and desert margins), no country is in a safe position over the century (*6*). Second, there is a growing number of adaptation responses emerging in both developed and developing countries (*7*), which make adaptation a challenge not only for the Global South. Third, the general understanding seems to be that the shaping and implementation of adaptation only come under national to local purview (*4*, *8*). This is too restrictive, as it does not account for risks from non- or maladaptation beyond national boundaries—on regional to global scales (*9*). Adaptation initiatives in one place may have adverse effects in neighboring places or interconnected ones, so that reducing vulnerability here can lead to increasing vulnerability there (*4*, *10*). One must also consider the risk that countries will not be able to adapt, which will have negative effects at regional to global scales (e.g., human migration, or slowdown in crop production).

Together, these arguments advocate for better consideration of trans-boundary effects of national adaptation strategies, and for strengthening bilateral to multilateral cooperation. Although the UNFCCC Cancún Adaptation Framework already stresses this point, it limits cooperation to countries that have common direct and revealed interests. Yet, the cascading effects of climate change impacts suggest that there will also be partly unpredictable ramified consequences. This emphasizes the need for the international community to anticipate impacts before they occur (i.e., address the unrevealed impacts) and consider possible indirect impacts (i.e., cascading effects).

Beyond simply providing funds for national-level adaptation, there is a need for enhancing a global sense of responsibility on the shaping of adaptation. This supposes the international community to improve the comprehensiveness of the existing Cancún Adaptation Framework. It could be inspired by the way the framework for mitigation has been progressively developed, i.e., definition of a common goal (the +2°C target established in Copenhagen in 2009) and references and tools [e.g., atmospheric pollution equivalent to one metric ton of CO_2 and Intended Nationally Determined Contributions (INDCs)] to track progress and efficiency.

EXPECTATIONS AND CHALLENGES. At least four major benefits are to be expected from a global adaptation framework. First, it would be a way to put nations of the world on the same road, as happened for mitigation. Second, it would provide incentives and guidance at the national level (*11*), which will stimulate design and implementation of adaptation strategies. Here again, the case of global mitigation is inspiring. Third, it would help address the underdebated question "Are we on track to adaptation?", which is complementary to "Are we on track to mitigation?", to decide whether the well below +2°C, if not +1.5°C, mitigation target established in Paris describes a sustainable future. Last, the better we track adaptation at the national level, the better we will be able to anticipate and avoid negative effects of non- or maladaptation on the regional to global scales.

Three major challenges arise and lay foundations for a post-2015 road map for climate negotiations on adaptation. First, we must define what a global adaptation goal should be, as this is the first building block of any tracking mechanisms. Both the Cancún Adaptation Framework and the Paris Agreement (*1*) remain imprecise on this, the latter referring to "enhancing adaptive capacity, strengthening resilience and

Institute for Sustainable Development and International Relations, Paris, France. Email: alexandre.magnan@iddri.org

reducing vulnerability to climate change, with a view to contributing to sustainable development." Such a multitarget perspective (i.e., adaptive capacity + resilience + vulnerability + sustainable development) is too broad and gives way to very intuitive and subjective interpretations. Something more specific is needed.

As food for thought, we draw on our previous work (*9*) in proposing the following definition of the global adaptation goal, which addresses a more precise issue (i.e., human security): the commitment of the international community to ensure human security in a "well below +2°C" world by the end of the century, meaning first, enhancing adaptation efforts when possible, and second, providing adequate answers for those whose security could not be covered in a well below +2°C world. We link the global adaptation goal to human security in response to widespread and crosscutting threats that can spread rapidly within and across nations and generate crises that challenge both governments and people. In addition, in line with the United Nations Office for the Coordination of Humanitarian Affairs, human security underscores the universality and interdependence of a set of freedoms that are fundamental to human life, as well as to societies' adaptive capacity to climate change (e.g., equity, access to safe environmental resources).

A second challenge calls on the scientific community to help move toward a more structured approach to adaptation and more explicit targets by defining criteria to capture adaptation national efforts. There is a long-standing and sensitive debate on indicators (*4*, *12*). On one hand, defining qualitative and/or quantitative metrics at the national level raises problems such as representativeness ("Do indicators capture what is really happening in the field?") and comparability (an indicator can be relevant for one country but not another). On "representativeness," for example, a national adaptation plan may not necessarily entail nationwide efficiency in the adaptation decision-making process. On "comparability," for example, an indicator reflecting the percentage change of the economic cost of extreme events in 2050 compared with that change in 1990 encounters the problem of discrepancies from one country to another in the extent of national databases and current levels of exposure. These limitations are inherent in the context-specific nature of adaptation.

On the other hand, as shown for mitigation within the UNFCCC context, a universal agreement requires negotiations to be based on a few clear criteria. Given that impacts are and will be worldwide, and non- or maladaptation has and will have transboundary effects, it is crucial to overcome the "intuitive and subjective" understanding of adaptation currently spread through negotiations. Even imperfect references to capture adaptation are needed to guide and delimit international discussions.

A way to reconcile the pros and cons of indicators could be for scientists to provide parties to the UNFCCC with an updated synthesis of benefits and limitations of existing methods to assess adaptation efforts qualitatively and quantitatively (*13*, *14*)—in line with the IPCC principle of being policy-relevant without being policy-prescriptive. Parties could discuss the relevance of those references from a policy point of view and identify indicators to apply at the country level, in accordance with national circumstances and country-driven principles enhanced in the Paris Agreement (Art. 2.2 and 7.5).

> ***"Even imperfect references to capture adaptation are needed to guide and delimit international discussions."***

One example comes from Mexico's INDC (*15*): reduce by 50% the number of municipalities considered "most vulnerable" to climate change. The key is to support knowledge coproduction (*16*) to define equilibrium between what is scientifically robust and what is politically acceptable, and eventually assess the feasibility of an indicator-based framework. This supposes that the scientific and the negotiation communities will make compromises. Scientists must accept an imperfect and rough estimate of adaptation efforts to be a foundation for international action. Parties must accept being challenged by scientists on the robustness of their criteria in order to avoid misinterpretations. For example, it would be crucial in the case of Mexico to clearly define what describes the "most vulnerable" municipalities.

Last, tracking adaptation efforts and transboundary negative consequences will raise political barriers. For example, some developing countries could be reluctant to report their adaptation efforts, depending on the way the international community will take them (e.g., encourage further efforts with more funding or prioritize countries showing less progress). Some developed countries may fear that their own authorities, populations, and stakeholders can blame them for insufficient nationwide efforts. Although it is difficult to envisage all political barriers now—particularly as they may be correlated to barriers inherent in the negotiation process on mitigation—it is important to pay attention to their emergence. Science has a vital role to play, notably, by demonstrating the usefulness of a global adaptation framework and by regularly bringing new empirical evidence on indirect and collateral effects of non- or maladaptation beyond national boundaries.

Three conditions will eventually determine whether Paris really laid foundations for a new era for climate change adaptation. The first is ratification of the Paris Agreement by April 2017 (*1*). The second is the ability of the international climate negotiation community to build a more comprehensive global adaptation framework and not uncouple mitigation and adaptation storylines over the 21st century. This will partly depend on the third condition: the effectiveness of the science-policy interface and the ability of the scientific community to help define practical criteria (i.e., specific adaptation goals reflecting national circumstances), design tracking protocols (i.e., how to aggregate national contributions and provide a global stocktaking), and develop research to assess adverse effects of non- and/or maladaptation (e.g., transdisciplinary analyses of concrete case studies). ■

REFERENCES

1. UNFCCC, The Paris Agreement (FCCC/CP/2015/L.9/Rev.1) (2015); http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf.
2. C. B. Field *et al.*, Eds., *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A, Global and Sectoral Aspects: Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, New York, 2014).
3. UNFCCC, Report of the Adaptation Committee to the Subsidiary Body for Scientific and Technological Advice (Forty-first session of COP20, Lima, Peru, FCCC/SB/2014/2) (2014).
4. I. Noble *et al.*, pp. 833–868 in (*2*).
5. A. K. Magnan, *Nature* **530**, 160 (2016).
6. V. Murray *et al.*, in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, C. B. Field *et al.*, Eds. (Cambridge Univ. Press, New York, 2012), pp. 487–542.
7. R. J. T. Klein *et al.*, pp. 899–943 in (*2*).
8. N. Mimura *et al.*, pp. 869–898 in (*2*).
9. A. Magnan, T. Ribera, S. Treyer, "National adaptation is also a global concern" (IDDRI, Working Paper No. 04/15, The Institute for Sustainable Development and International Relations, Paris, 2015).
10. S. Juhola, E. Glaas, B.-O. Linner, T.-S. Neset, *Environ. Sci. Policy* **55**, 135 (2016).
11. J. Helgeson, J. Ellis, "The role of the 2015 agreement in enhancing adaptation to climate change" (OECD, Climate Change Expert Group, Paper No. 2015-1, Organization for Economic Co-operation and Development, Paris, 2015).
12. J. Hinkel, *Glob. Environ. Change* **21**, 198 (2011).
13. A. Lesnikowski, J. D. Ford, R. Biesbroek, L. Berrang-Ford, S. J. Heymann, *Nat. Clim. Change* **6**, 261 (2015).
14. J. D. Ford *et al.*, *Nat. Clim. Change* **5**, 967 (2015).
15. México Gobierno de la Repùblica, INDC (Government of Mexico, 2015); http://bit.ly/MexicoINDC.
16. N. L. Klenk *et al.*, *Science* **350**, 743 (2015).

10.1126/science.aaf5002

BOOKS *et al.*

The entirety of Gates of the Arctic National Park in Alaska lies north of the Arctic Circle.

PHOTO: PAXSON WOELBER/WIKIMEDIA COMMONS

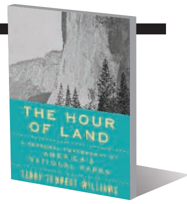
NATIONAL PARKS

Meditations on conservation

An environmental activist urges a renewal of the American national park idea

Reviewed by **Jared Farmer**

The Hour of Land
A Personal Topography of America's National Parks
Terry Tempest Williams
Sarah Crichton Books, 2016. 416 pp.



The centennial of the U.S. National Park Service has thus far been muted, in no small part because the anniversary year has coincided with a noisy presidential election. News about the agency has not helped: budget shortfalls and scandals, too. And yet the idea of the park system remains beloved, and the flagship parks are being "loved to death" more than ever.

It would have been easy for Terry Tempest Williams to fall back on "the best idea we ever had"—an interpretation articulated by Wallace Stegner in 1983 and popularized by Ken Burns in 2009. Instead, Williams asks hard questions about the current relevance and original goodness of America's parks. She offers a poetic revision to the Organic Act of 1916, which mandated the conservation of scenery and wildlife for the enjoyment of the public in such a manner as to leave them unimpaired. In her 400-page mission statement, Williams updates "enjoyment" to spiritual renewal, specifies that "the public" means more than white people, and insists that "unimpaired" means what it says.

To write the 12 chapters that compose *The Hour of Land*, Williams pilgrimaged to 12 units in the park system. She chose her itinerary with care, mixing the obscure and the famous, ranging from Acadia to Alcatraz, the Gates of the Arctic to the Gulf Islands. The varied locations inspired various genres, including prose poetry, criticism, personal correspondence, and reportage.

Like Williams's oeuvre, the entire book is, on some level, a memoir. Having written so affectively about the women in her life, Williams here gives more attention to men: her husband, her estranged brother, her adopted Rwandan son, her climate activist ally Tim DeChristopher, and especially her Mormon father, a retired oil pipe layer. John Henry Tempest loves Grand Teton National Park as much as the Rockefellers—the other family that appears prominently in these pages—but he also shoots prairie dogs for fun and feels "proud of the scars I've left in the West" with his family-owned Tempest Company.

"Strange things happen when Terry's around," her dad confirms. In *The Hour of Land*, Williams attracts rare woodpeckers, meets an upright grizzly, witnesses a horizontal rainbow, and escapes death by wildfire.

Although she has moved beyond the faith of her father, Williams speaks in religious terms. Enraptured and enraged by our world, she blesses it with holy words like care, ceremony, compassion, humility, inclusion, integrity, restraint, and reverence.

For Williams, the personal is the spiritual is the political. Throughout her centrifugal text, she returns to two controversies. The first is the U.S. oil-and-gas boom of the Bush-Obama years. She visits park beaches in Florida befouled by Deepwater Horizon, park vistas in North Dakota spoiled by fracking rigs, and potential parkland in Utah trampled and fragmented by drill pads. She turns her field observations into calls for societal divestment from fossil fuel.

Her second quarrel is with the "land transfer movement" in the mountain West. From renegade rancher Cliven Bundy to Utah representative Rob Bishop (who recently let the Land and Water Conservation Fund expire), many influential figures in the region have demanded that the U.S. government "give back" federal lands. With her chapter on Gettysburg National Military Park, Williams implicitly compares the western "Sagebrush Rebellion" to the southern "Lost Cause"—a pair of states' rights movements fueled by the resentments of gun-loving white men.

Williams shows more sympathy for Native peoples who were dispossessed from wilderness parks such as Yellowstone. To have an environmentalist of her stature address this violent history is a sign that the 1990s rift between environmental historians and conservation biologists (the "wilderness debate") has largely healed.

Instead of telling simple stories of innocent preservationists saving pristine places, Williams relates histories of conflict, struggle, money, and power. "We, the people, have made mistakes," she summarizes.

In *The Hour of Land*, reconciliation follows truth. The author sees hope in the Blackfeet Nation's demands for comanagement of Glacier National Park. She takes pride in the first black president's authorization of a national monument in honor of César Chávez. And she beseeches Barack Obama, in his final months in office, to go further.

At a recent federal auction, the author purchased drilling rights to lands near her home in Utah. To avoid violating the law, she founded Tempest Exploration Company, LLC. Writing for the *New York Times*, Williams announced her company's intent to produce "energy" to "fuel moral imagination" (*1*). On Facebook, Williams added that her father accepted the call to serve as chairman of the board but only after "serious and soulful conversations." This episode—environmental activism meets performance art meets family drama—serves as a fitting introduction to a sincerely disobedient book.

The reviewer is at the Department of History, State University of New York, Stony Brook, NY 11794, USA. Email: jared.farmer@stonybrook.edu

REFERENCES

1. T. T. Williams, "Keeping my fossil fuel in the ground," *The New York Times*, 29 March 2016, p. A25.

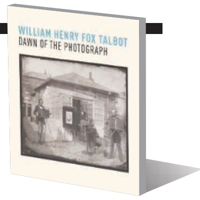
10.1126/science.aaf7129

PHOTOGRAPHY

Candid camera

A new exhibition provides an intimate look at a pioneer of modern photography

By **Andrew Robinson**



William Henry Fox Talbot
Dawn of the Photograph
Russell Roberts and Greg Hobson
Scala Arts & Heritage Publishers, 2016. 176 pp.

Fox Talbot: Dawn of the Photograph
Science Museum, London
Through 11 September 2016

The credit for inventing photography is complicated and contested. Nicéphore Niépce, Louis Daguerre, Henry Fox Talbot, and the scientific polymath John Herschel each have a claim. The term itself (from the Greek for "drawing with light") seems to have been coined in 1839 by Herschel. However, there can be little doubt that Fox Talbot—the subject of a new exhibition at London's Science Museum—was the key pioneer.

In 1835, Fox Talbot produced the first photograph on paper: an image showing a latticed window of his country home at Lacock Abbey, which he termed a "photogenic drawing." Then, in 1840–1841, he discovered and patented a process for producing negatives from which he could make prints, which he termed a "calotype" (Greek for "beautiful impressions"). Various improvements to this basic method, beginning with the invention of the wet collodion process and the adoption of a glass negative invented by Frederick Scott Archer, would dominate the field until the arrival of commercial digital photography in the 1990s.

Fox Talbot: Dawn of the Photograph is, surprisingly, the first major London exhibition about the British inventor. A thought-provoking mixture of technology and art, the exhibition displays numerous images taken by Fox Talbot and several contemporary photographers who adopted his calotype process. The photos range from portraits of family members to a fascinating shot from 1844 of Nelson's Column, then under construction in London's Trafalgar Square, in which the quotidian words "NO BILLS TO BE STUCK" are plainly visible near the base. About 100 of these images appear in the well-produced and informative companion book, *William Henry Fox Talbot: Dawn of the Photograph*, compiled by the exhibition's cocurators, Greg Hobson and Russell Roberts.

Nelson's Column under construction, Trafalgar Square, London, April 1844.

Fox Talbot revealed his photogenic drawings in public at London's Royal Institution in 1839, in a rapid response to the debut of the "daguerreotype"—a photographic image made on a copper plate—in Paris less than 3 weeks earlier. They immediately intrigued Michael Faraday. "No human hand has hitherto traced such lines as these drawings display; and what man may hereafter do, now that Dame Nature has become his drawing mistress, it is impossible to predict," he presciently remarked. Echoing Faraday's pronouncement, Fox Talbot entitled the earliest book to be illustrated by actual photographic prints *The Pencil of Nature* when he published its initial volume in 1844. In a note to the reader, he emphasized that the plates had been "impressed by the agency of Light alone, without any aid whatever from the artist's pencil." He called them "sun pictures" rather than "engravings in imitation," as some people had thought them to be. The new name was better suited, given the astonishing immediacy of certain of the images. One, which showed lace, famously fooled the photographer's friends into thinking that they were looking at the real thing.

Daguerre was a considerable painter and was led to the daguerreotype by his interest in dioramas. By contrast, despite publishing with distinction in fields ranging from mathematics to Assyriology, Fox Talbot was keenly aware that he possessed no artistic ability. In fact, he was led to photogenic drawing precisely as a result of his frustration with his lack of artistic skill. As he noted in 1839, the first inkling of photography struck him while visiting Lake Como in Italy on his honeymoon in 1833. Using a camera lucida, a tricky drawing aid that allowed an artist to see both the subject and the paper simultaneously, he attempted to draw the view from his villa. His tracing was "melancholy to behold," as can be verified from its appearance in the exhibition. This failure, and the fleeting images Fox Talbot had seen using another aid to drawing, a camera obscura, provoked his keen desire to "cause these natural images to imprint themselves durably, and remain fixed upon the paper!"

Daguerreotypes enjoyed a boom in the 1840s due to the amazing clarity of their images. One of the earliest to survive, "Les Coquillages" (shells), taken in 1839, has been borrowed from France for the exhibition. Others on display include portraits of Faraday and Fox Talbot. But each daguerreotype was unique, whereas the calotype process permitted multiple reproductions.

After London's Great Exhibition in 1851, an estimated 20,000 calotype prints were made for the 137 presentation sets of the show's jury reports (also on display). The prints included images such as "Steam Engine," "Large Anchor," "Glass Fountain," "Collection of Feathers," "Nymph Preparing for the Bath," and "Cholera"—scenes that depict a panorama of Victorian British life and industry.

Today, anyone can take color photographs with minimum effort and expense. After viewing the devoted labors of the pioneers in *Dawn of the Photograph*, it is hard to avoid feeling that we have both lost something and gained enormously from the total democratization of photography.

The reviewer is the author of The Story of Measurement *(Thames & Hudson, London, 2007). Email: andrew@andrew-robinson.org*

10.1126/science.aaf8965

LETTERS

Edited by **Jennifer Sills**

Speaking out against blood antiquities

THE DESTRUCTION OF Iraq and Syria's heritage has unfolded on an "industrial scale," according to UNESCO, the United Nation's cultural agency (*1*). In recent years, we have watched helplessly as terrorist organizations looted museums and thousands of archaeological sites and then sold ancient treasures to willing buyers in the United States, Europe, and Asia (*2*). After oil, these "blood antiquities" may be the Islamic State's second largest source of revenue (*3*). The Middle East's enduring past has become one more tragic casualty of war.

At last, we do not have to be silent witnesses to this carnage. In a rare moment of bipartisan accord, the U.S. Congress passed the Protect and Preserve International Cultural Property Act. Signed into law by President Obama on 9 May (*4*), the act will delimit the illicit antiquities market by, for example, establishing an interagency coordinating committee and imposing import restrictions on cultural materials from Syria.

Far more must be done. In a recent report, preservation advocates and legal experts (*5*) suggest that the U.S. Justice Department should appoint more prosecutors with expertise in heritage crimes, and that museums must become more transparent about the ownership of antiquities in their care. A recent letter by 11 major archaeological and museum organizations has also pressed the U.S. Congress to resume paying its annual dues to UNESCO (*6*). As one of the world's largest consumers of antiquities, the United States should continue to take action that will curtail the dark trade in blood antiquities.

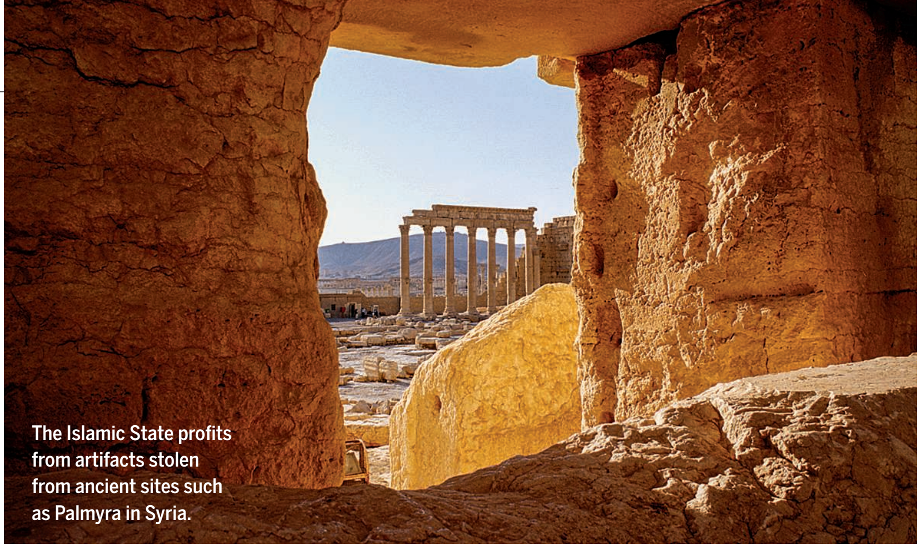
Chip Colwell

Department of Anthropology, Denver Museum of Nature and Science, Denver, CO 80205, USA. Email: chip.colwell@dmns.org

REFERENCES

1. A. Osborn, "Islamic State looting Syrian, Iraqi sites on industrial scale –UNESCO," *Reuters* (2015); http://uk.reuters.com/article/uk-mideast-crisis-unesco-idUKKCN0PC1OS20150702.
2. S. Mulder, "The blood antiquities funding ISIL," *Al Jazeera* (2014); www.aljazeera.com/indepth/opinion/2014/11/blood-antiquities-funding-isil-2014119113948461658.html.
3. "Looted in Syria—and sold in London: The British antiques shops dealing in artifacts smuggled by ISIS," *The Guardian* (2015); www.theguardian.com/world/2015/jul/03/antiquities-looted-by-isis-end-up-in-london-shops.
4. H.R.1493–Protect and Preserve International Cultural Property Act (www.congress.gov/bill/114th-congress/house-bill/1493).
5. "#CultureUnderThreat: Recommendations for the U.S. Government" (2016); http://taskforce.theantiquitiescoalition.org/wp-content/uploads/2015/01/Task-Force-Report-April-2016-Complete-Report.pdf.
6. Society of American Archaeologists, World Heritage Letter (www.saa.org/Portals/0/SAA/GovernmentAffairs/WorldHeritageLetter_12_10_2015_Final.pdf).

10.1126/science.aaf9132



The Islamic State profits from artifacts stolen from ancient sites such as Palmyra in Syria.

PHOTO: DOMINIK PABIS/ISTOCKPHOTO.COM

Institutionalizing creationism

BIOLOGY FACULTY WHO teach evolution at U.S. colleges and universities often worry about the efforts of creationists to include the teaching of "intelligent design" in publicly funded high school biology courses. Now we also have cause to worry about students at publicly funded colleges and universities earning science credits for learning creationism.

The Western Interstate Commission for Higher Education (WICHE) is developing an Interstate Passport Initiative, funded in part by the U.S. Department of Education, which would streamline the learning outcomes for courses across institutions to facilitate the transfer of credits (*1*). Unfortunately, with the Passport Initiative, WICHE proposes making the creationist "teach the controversy" strategy as a standard part of college biology courses. In their document "Faculty handbook: Constructing your institution's Passport block," WICHE suggests that to demonstrate scientific literacy, students should "watch the Ken Hamm [sic]–Bill Nye evolution-creation debate and evaluate the scientific evidence and arguments used by the participants" (*2*).

This suggestion validates creationism as science by stating explicitly that both participants have scientific evidence. Middle school, high school, and college instructors who support creationism can point to the WICHE Passport Initiative as evidence that there is a scientific debate that includes creationism. The Answers in Genesis website has already promoted the debate as a way to get creationism into science classrooms (*3*).

If the goal of the curriculum is to help students use scientific evidence to debunk myths, the suggested class activity should be rephrased to read, "Watch the Ken Ham–Bill Nye evolution-creation debate and evaluate the arguments used by the participants." However, even with better wording, by including the debate in a science class, WICHE is promoting the use of the Ham-Nye debate as an example of a scientific controversy. There are hundreds of genuine biological debates, both current and historical, that good educators can make interesting. WICHE should choose real examples of scientific debates and avoid advocating for creationism in science classrooms.

A student who takes general education courses at a WICHE Passport institution will soon be able to transfer the credits to any other Passport institution. The receiving institution cannot reject individual courses from approved institutions. Currently, WICHE lists 24 public institutions representing more than 150 campuses in seven U.S. states as participants in developing the Passport Initiative. WICHE plans to expand the Passport Initiative to six more states. As the Initiative grows, more and more public postsecondary institutions will be awarding science credits for courses that include creationism. To prevent the insertion of religion into science classrooms, scientists must speak out against the Passport Initiative until WICHE removes creationism from their suggested curriculum.

Michael Baltzley

Department of Biology, Western Oregon University, Monmouth, OR 97361, USA. Email: baltzlem@wou.edu

REFERENCES

1. Western Interstate Commission for Higher Education, The Interstate Passport (www.wiche.edu/passport/home).
2. Interstate Passport, "Faculty handbook: Constructing your institution's Passport block" (2016); p. 43; www.wiche.edu/files/info/Faculty%20HB-CYIPB%20March%202016%20%28REV%204-08-16%29.pdf.
3. Answers in Genesis, "Public schools and the

Bill Nye/Ken Ham debate" (2014); https://answersingenesis.org/creation-vs-evolution/public-schools-and-the-bill-nye-ken-ham-debate/.

10.1126/science.aaf7386

The language within

IN THE NEWS In Depth story "How sign languages evolve" (22 April, p. 392), C. Matacic suggests that recent evidence on the emergence of novel sign languages offers clues to the evolution of linguistic complexity. She reports that the results "show that social interaction is essential for language evolution." Although establishing a fully developed sign language in a few generations is remarkable, this phenomenon does not shed light on how human language evolved.

All of the sign languages discussed in the News story developed in individuals who had modern brains with completely evolved internal computational systems for language with its full syntactic complexity. Rather than demonstrating linguistic evolution, these observations elucidate how our pre-existing internal language system converts internal linguistic forms into speech or sign—a process known as externalization. Evidence of the common convergence to similar external signed language outcomes over just a few generations in different populations reinforces the single most fundamental finding of modern linguistic theory: that there is a common genomic underpinning to the computational syntactic "engine" in all humans (*1*, *2*).

Language as an internal cognitive system must be distinguished from its externalization as speech or sign. Internal language involves a wide array of hierarchically structured expressions that describe complex concepts and intentions (*1*, *2*). To convert concepts into speech, the syntactic hierarchical structure must be "flattened" into a linear sequence of wordlike elements (*3*). The externalization process, whether the words take the form of speech or sign, is a secondary process (*2*) that evolved more recently than internal cognition (*3*–*5*).

Marinus A. C. Huybregts,[1] Robert C. Berwick,[2] Johan J. Bolhuis[3,4]*

[1]Utrecht University, 3512 JK, Utrecht, Netherlands. [2]Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Utrecht University, 3584 CH, Utrecht, Netherlands. [4]Department of Zoology and St. Catharine's College, University of Cambridge, Cambridge, CB2 3EJ, UK.

*Corresponding author. Email: j.j.bolhuis@uu.nl

REFERENCES

1. R. C. Berwick, A. D. Friederici, N. Chomsky, J. J. Bolhuis, *Trends Cogn. Sci.* **17**, 89 (2013).
2. M. B. H. Everaert *et al.*, *Trends Cogn. Sci.* **19**, 729 (2015).
3. R. C. Berwick, N. Chomsky, *Why Only Us: Language and Evolution* (MIT Press, 2016).
4. J. J. Bolhuis, I. Tattersall, N. Chomsky, R. C. Berwick, *PLOS Biol.* **12**, e1001934 (2014).
5. J. J. Bolhuis, M. Everaert, Eds., *Birdsong, Speech & Language: Exploring the Evolution of Mind and Brain* (MIT Press, 2013).

10.1126/science.aag0933

ERRATA

Erratum for the Report "A plant miRNA contributes to antibacterial resistance by repressing auxin signaling" by L. Navarro *et al.*, *Science* **351**, aae0382 (2016). Published online 22 January 2016; 10.1126/science.aae0382

Erratum for the Report "Small RNA duplexes function as mobile silencing signals between plant cells" by P. Dunoyer *et al.*, *Science* **351**, aae0387 (2016). Published online 22 January 2016; 10.1126/science.aae0387

Erratum for the Research Article "Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense" by A. Deleris *et al.*, *Science* **351**, aaf2336 (2016). Published online 22 January 2016; 10.1126/science.aaf2336

REVIEW SUMMARY

TRANSCRIPTION

Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut

Nick J. Proudfoot*

BACKGROUND: Genes correspond to single transcription units, starting from the promoter and ending at the terminator. Terminating gene transcription is directly coupled to mRNA processing, which occurs cotranscriptionally. When RNA polymerase II (Pol II) reaches the gene end, it first slows down over the terminator. This is partly because 3′-end cleavage and polyadenylation (CPA) complex is recruited onto Pol II when poly(A) signals appear in the nascent transcript. This nascent transcript will often invade the DNA duplex to form an R-loop structure, which induces further polymerase slowdown. During this time, CPA releases mRNA from chromatin into eventual cytoplasmic translation. Pol II continues to transcribe its DNA template after mRNA release. However, this is short-lived, as an exonuclease (Xrn2) degrades the transcript from its 5′ end. When this molecular torpedo catches up with Pol II, then conformational shockwaves are transmitted into its active site, which releases Pol II from the DNA template. Pol II is then free to restart transcription on another gene promoter.

ADVANCES: The above process of Pol II termination appears surprisingly complex, but provides unanticipated layers of gene regulation. First, most protein-coding genes generate multiple mRNAs of different lengths caused by the use of alternative poly(A) sites (APA), which in turn is dictated by gene termination. Alternative mRNAs with shorter or longer 3′-untranslated region (3′UTR) sequences possess different sequence codes for how long to survive or where in the cell to translate their proteins. Second, termination of transcription is employed as a quality-control mechanism. Transcription errors occur either because the DNA template is damaged or because the RNA is mis-synthesized and induce premature termination before reaching the gene end. These truncated transcripts are rapidly degraded and, if Pol II becomes arrested on the DNA template, then it is degraded in situ by the proteasome to allow subsequent rounds of transcription.

Recent studies reveal that cellular stress such as osmotic or heat shock, as well as viral infection or cancer-inducing mutations, can all promote aberrant termination. Under these varied conditions, many genes fail to terminate transcription. The resulting extensive readthrough transcription can cause massive deregulation of downstream gene expression.

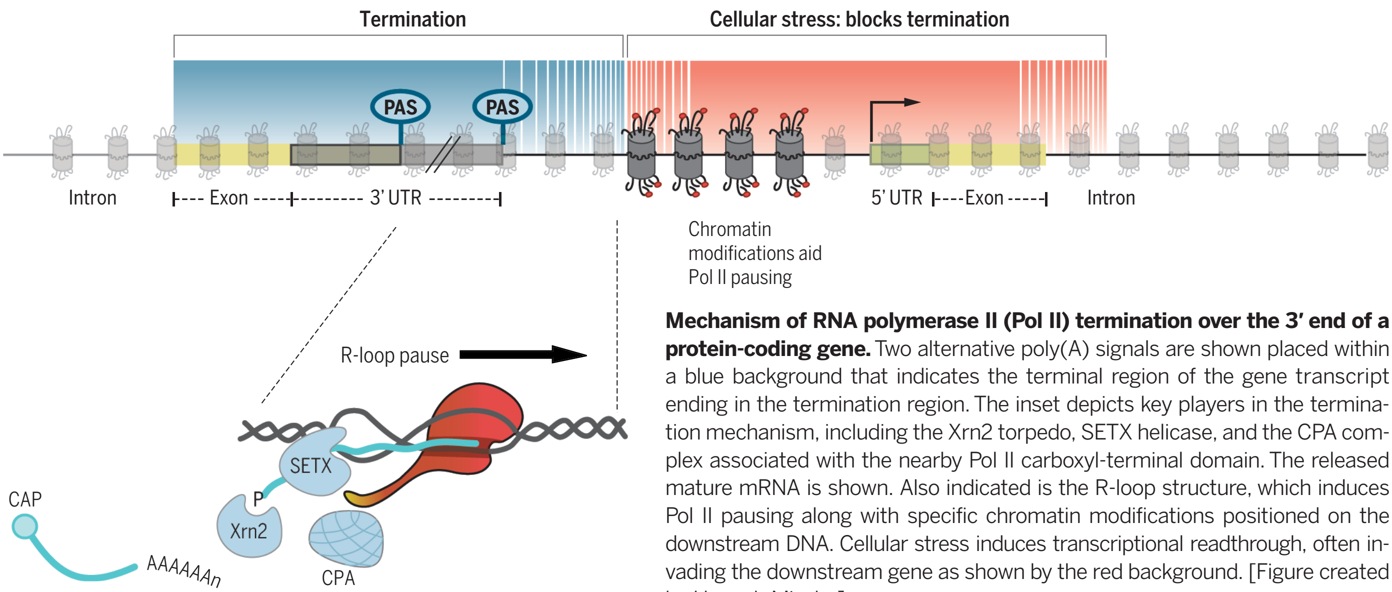
OUTLOOK: Many questions remain about the mechanism and regulation of transcriptional termination. Exactly how degradation of the transcript by Xrn2 together with CPA and various helicases promotes Pol II termination remains poorly understood. Structural changes occurring within Pol II to promote this effect are currently unknown. Their resolution will need new technology, such as cryo–electron microscopy.

The regulation of APA to generate mRNA with different 3′UTRs is similarly poorly understood. Although changes in CPA factor levels or in the transcription process itself can affect APA, dominant factors and mechanisms used in biology to achieve this regulation remain enigmatic. The easy perturbation of termination resulting in readthrough transcripts appears to be at odds with the elaborate mechanisms in place to stop the Pol II juggernaut. It is evident that the field of Pol II termination has many surprises in store for future research into this fascinating process. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aad9926

The list of author affiliations is available in the full article online.
*Corresponding author. Email: nicholas.proudfoot@path.ox.ac.uk
Cite this article as N. J. Proudfoot, *Science* 352, aad9926 (2016). DOI: 10.1126/science. aad9926.



Mechanism of RNA polymerase II (Pol II) termination over the 3′ end of a protein-coding gene. Two alternative poly(A) signals are shown placed within a blue background that indicates the terminal region of the gene transcript ending in the termination region. The inset depicts key players in the termination mechanism, including the Xrn2 torpedo, SETX helicase, and the CPA complex associated with the nearby Pol II carboxyl-terminal domain. The released mature mRNA is shown. Also indicated is the R-loop structure, which induces Pol II pausing along with specific chromatin modifications positioned on the downstream DNA. Cellular stress induces transcriptional readthrough, often invading the downstream gene as shown by the red background. [Figure created by Hannah Mischo]

REVIEW

TRANSCRIPTION

Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut

Nick J. Proudfoot*

Terminating transcription is a highly intricate process for mammalian protein-coding genes. First, the chromatin template slows down transcription at the gene end. Then, the transcript is cleaved at the poly(A) signal to release the messenger RNA. The remaining transcript is selectively unraveled and degraded. This induces critical conformational changes in the heart of the enzyme that trigger termination. Termination can also occur at variable positions along the gene and so prevent aberrant transcript formation or intentionally make different transcripts. These may form multiple messenger RNAs with altered regulatory properties or encode different proteins. Finally, termination can be perturbed to achieve particular cellular needs or blocked in cancer or virally infected cells. In such cases, failure to terminate transcription can spell disaster for the cell.

Genes are defined as regions of the genome that correspond to a single transcription unit (TU), starting from the promoter and ending at the terminator. Although promoters are often well characterized, less is known about the mechanism and regulation of transcriptional termination.

Prokaryotes versus eukaryotes

For prokaryotic genes, protein expression units (cistrons) are usually clustered into tandem arrays transcribed as a single TU, creating a polycistronic messenger RNA (mRNA). Failure to terminate transcription results in the inclusion of extra cistrons in the extended mRNA that may cause the production of unwanted proteins with adverse biological consequences (*1*). The basic mechanism of termination in *Escherichia coli* is well defined. Formation of an RNA hairpin structure, immediately followed by an oligo(U) sequence in the nascent transcript, triggers termination (*2*). Alternatively, the adenosine 5´-triphosphate (ATP)–dependent translocase Rho can promote termination by recognizing a loosely defined C-rich sequence (Rho utilization transcript, RUT) (*3*). After initial polymerase binding, hexameric Rho translocates and unravels the nascent RNA in association with the elongating polymerase (*4*). Contacts between an RNA hairpin or Rho and the polymerase somehow trigger conformational changes that switch the polymerase's enzymatic mode from elongation to termination. In prokaryotes, mRNA translation occurs on transcripts still being made by RNA polymerase (cotranscriptional). Translation elongation along the mRNA template can remove RNA hairpin structures or block access of Rho to RUT sites. Either way, translation can directly regulate termination and the consequent extent of TUs (*5*).

Eukaryotic gene transcription is fundamentally different from that of prokaryotes, as it occurs in the nucleus, separate from the cytoplasmic translation apparatus. Furthermore, eukaryotes employ three different classes of RNA polymerase (Pol). Pol II transcribes all protein-coding genes to generate mRNA, as well as many noncoding RNAs (ncRNAs). ncRNA can either be abundant and stable, such as small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA), or be present at low levels and rapidly degraded, such as long noncoding RNA (lncRNA) that may run between or overlap with protein-coding genes (*6*). Pol I transcribes the highly abundant ribosomal RNA (rRNA) precursor, which is cotranscriptionally processed to mature 28*S*, 18*S*, and 5.8*S* rRNA, whereas Pol III transcribes transfer RNA (tRNA) and 5*S* rRNA. All eukaryotic mRNAs are monocistronic, with a short RNA tract before and a longer one after the coding region (5′ and 3′ untranslated regions, or UTRs). The 5′UTR begins with a 5′ terminal Cap structure, whereas the 3′ UTR ends with a polyadenylate [poly(A)] tail. Both these terminal mRNA modifications are formed as part of pre-mRNA processing that occurs cotranscriptionally and is also coordinated with removal (splicing) of introns that separate the coding exons. These complex RNA processing reactions are all required to generate translatable mRNA, which is then exported through the nuclear pore to sites of cytoplasmic translation.

Failure to terminate transcription in eukaryotic genes may have severe consequences for gene expression. For protein-coding genes arranged in tandem, readthrough transcripts from a nonterminated upstream gene will run into the promoter of the downstream gene and restrict its activity by a process called transcriptional interference (*7*, *8*). This will in turn prevent Cap addition to the downstream gene transcript, as this can only occur on a triphosphorylated 5′ end. For genes arranged in convergent orientation, termination defects may result in the formation of overlapping transcripts that down-regulate gene expression by triggering RNA interference (RNAi) pathways (*9*). In severe cases, failure of convergent genes to terminate transcription will result in molecular collision between Pol II transcribing opposite DNA template strands (*10*, *11*). Failed termination may also result in Pol II elongation complexes running into regions of the genome undergoing DNA replication. Collision with DNA polymerase complexes may disrupt DNA synthesis and trigger DNA damage and genome instability (*12*). The extensive lncRNA transcriptome increases the likelihood of potential interference problems between TUs. Failure of lncRNA to terminate transcription may also cause interference with adjacent protein-coding genes (*13*), even though our current understanding of lncRNA termination is rudimentary.

The advent of high-throughput sequencing has made it possible to visualize nascent transcription with a variety of techniques (*14*, *15*). Consequently, it is now possible to directly map the extent of transcription past the poly(A) site (PAS) of a gene. Many Pol II genes display gradual termination profiles across multiple kilobases, whereas others terminate abruptly soon after the PAS. Here, I describe current understanding of how RNA Pol II terminates transcription, mainly focusing on mammalian protein-coding genes, but also with reference to other eukaryotic systems that exemplify specific features. I will start with a consideration of how the chromatin template signals Pol II to either slow down (pause) or completely stop (arrest). I will then consider how transcript processing and degradation can trigger Pol II termination. Finally, I will describe how termination can often be modulated to allow enhanced gene regulation or perturbed to cause genetic disease.

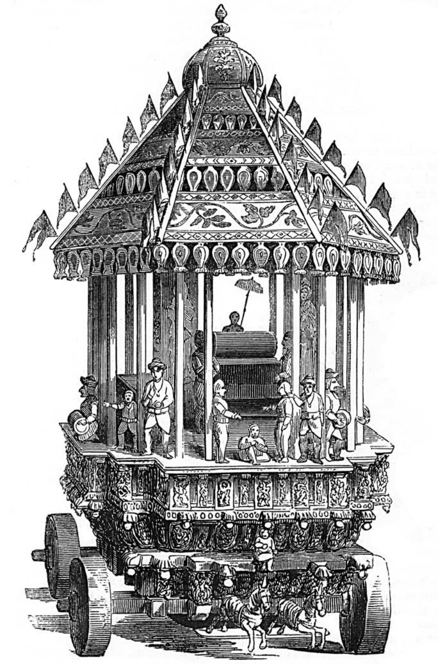
Transcriptional pausing

Pol II is uniquely endowed with an extra protein segment separate from the main globular enzyme that derives from the carboxyl-terminal domain (CTD) of Rpb1 (*16*). The CTD plays a critical role in coordinating cotranscriptional RNA processing: capping, splicing, and 3′-end cleavage and polyadenylation. In mammals, it comprises a relatively unstructured polypeptide of 52 heptad repeats (consensus Tyr^1-Ser^2-Pro^3-Thr^4-Ser^5-Pro^6-Ser^7) that are rich in differentially phosphorylated serine residues. In particular, phospho-Ser^2 is associated with gene 3′ ends and interacts with a large complex of cleavage and polyadenylation (CPA) factors that generate mRNA 3′ ends (*17*–*19*). A further important difference between Pol II and all other types of RNA polymerase transcription is that Pol II–transcribed genes in mammals vary in length, from a few hundred nucleotides for snRNA genes through to protein-coding genes that may exceed 100 kb in length. It is evident that Pol II must have the capacity to be highly processive; thus, termination mechanisms need to be sufficiently robust to stop this molecular juggernaut (Box 1).

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK.
*Corresponding author. Email: nicholas.proudfoot@path.ox.ac.uk

Box. 1.

Juggernaut was the name given for a huge wagon bearing an image of the god Krishna drawn annually in procession at Puri in Orissa, India, from the 1600s. Some devotees were crushed under its wheels in sacrifice. It now means a massive, unstoppable vehicle.



CAR OF JUGGERNAUT.

Pol II is capable of directly sensing its passage across a functional PAS. Depletion of key components of CPA, such as cleavage-polyadenylation specificity factor–73 (CPSF-73) and cleavage stimulatory factor–64 (CstF-64), reduces this Pol II pausing effect (*15*). Apparently, CPA recruitment to the Pol II elongation complex as it traverses the PAS induces an appreciable slow-down effect on Pol II elongation. Furthermore, in vitro transcription experiments indicate that the interaction of CPSF and CstF components with Pol II transcribing through a gene PAS can have marked pausing effects on transcription that result in the gradual release of Pol II from the DNA template (*20*). This termination process apparently occurs independently of PAS cleavage, arguing that Pol II conformational changes alone can induce substantial levels of transcriptional termination. Other features of the chromatin template may also induce pausing and in turn increase the dwell time of Pol II over the PAS. This will enhance CPA association with the PAS and the consequent 3′-end processing and transcriptional termination. Perhaps the most common type of transcriptional pausing for Pol II is caused by chromatin structure, especially the core nucleosome. Although it is clear that Pol II transcribes nucleosomal templates, it is also the case that nucleosome-free or nucleosome-depleted templates are more readily transcribed.

Pol II pausing can also be induced by hybridization of the nascent transcript with the antisense DNA strand outside the elongation complex. This results in the formation of RNA:DNA hybrids and displacement of the sense DNA strand, a structure referred to as an R loop (*21*, *22*). R-loop formation is favored by the act of Pol II transcription, because the DNA template behind the elongation complex is depleted in nucleosomes that are transiently displaced during the transcription process. The DNA double helix is also torsionally underwound (negatively supercoiled). Once formed, R loops can persist, especially as the RNA:DNA hybrid is thermodynamically more stable than duplex DNA. Also, they are often associated with G-rich regions of transcribed genes because the displaced sense DNA can form stabilizing G quadruplex structures (*22*). R loops were originally observed in budding yeast when pre-mRNA packaging by the THO complex was inactivated by gene deletion (*23*). Similarly, defective splicing can lead to enhanced R-loop formation (*24*). In both cases, accumulation of nascent RNA in close proximity to the underwound, just transcribed DNA template leads to RNA:DNA hybrid formation.

R-loop accumulation induced by mRNA packaging or splicing defects results in increased DNA damage. This is due to the mutagenic nature of the single-strand template DNA in the R loop, which leads to single- and then double-strand breaks with elevated levels of DNA recombination (*21*). Helicases such as Sen1 in yeast can act to remove these potentially harmful structures. Thus, loss of Sen1 gives an equally severe DNA damage phenotype to loss of THO (*25*). In mammalian cells, the homolog of Sen1, called Senataxin (mutant genes cause various neurological diseases; for example, ataxia oculomotor apraxia type 2 and amyotrophic lateral sclerosis type 4), is similarly required to resolve R loops but also plays a direct role in promoting more efficient termination (*26*). The recruitment of Senataxin to terminator regions is likely mediated by the creation of a specific Pol II CTD mark on an arginine residue present at the 7th position of the variant 31st heptad repeat (Arg^{1810}). Symmetric dimethylation of this residue by the methyltransferase PRMT5 recruits first SMN (survival of motor neuron disease associated protein), which then recruits Senataxin. Loss of any of these factors causes an accumulation of R loops and a defect in Pol II termination (*27*). An alternative Senataxin recruitment pathway involves the DNA repair factor BRCA1. This is recruited to R loops especially at Pol II terminators and in turn directly recruits Senataxin to effect the rapid resolution of R loops, promoting Pol II termination and preventing DNA damage (*28*).

Although R-loop structures display an intrinsic slow-down effect on Pol II elongation, they have also been shown to induce low-level antisense transcription. This may result in the formation of transient double-stranded RNA (dsRNA), which will in turn trigger an RNA interference effect mediated by nuclear Dicer and Ago proteins. This leads to dimethylation of histone H3K9 by the histone methyltransferase enzyme G9a-GLP (G9a-like protein) and consequent recruitment of HP1γ (heterochromatin protein 1γ), effectively creating localized patches of repressed chromatin (*29*). These will act to perpetuate and enhance Pol II pausing over R-loop–associated termination regions and are a feature of relatively short and ubiquitously expressed genes (Fig. 1A depicts different types of Pol II pausing).

Transcriptional arrest

This involves the irreversible association of Pol II with the DNA template so that it cannot be displaced by PAS-mediated termination mechanisms to allow efficient recycling. Instead, it is targeted for proteolytic degradation (Fig. 1B). PAS-dependent termination may be viewed as a productive mechanism allowing reuse of Pol II, whereas Pol II arrest can be viewed as nonproductive because Pol II is degraded on the DNA template.

The chemical modification or damage of DNA by oxidation or ultraviolet treatment results in arrested Pol II transcription complexes. Ubiquitin ligases are recruited to such complexes, resulting in Pol II degradation, effectively clearing the DNA template to allow new rounds of transcription (*30*). R loops may also restrict the passage of the replication fork formed in the wake of DNA replication (*31*). Such collisions between replication and stalled transcription complexes underlie fragile sites in the genome (*32*). Some trypanosome species convert the DNA base T to glucosyl hydroxymethyl uracil (base J) (*33*). Base J is found at the end of many TUs, especially those arranged in convergent

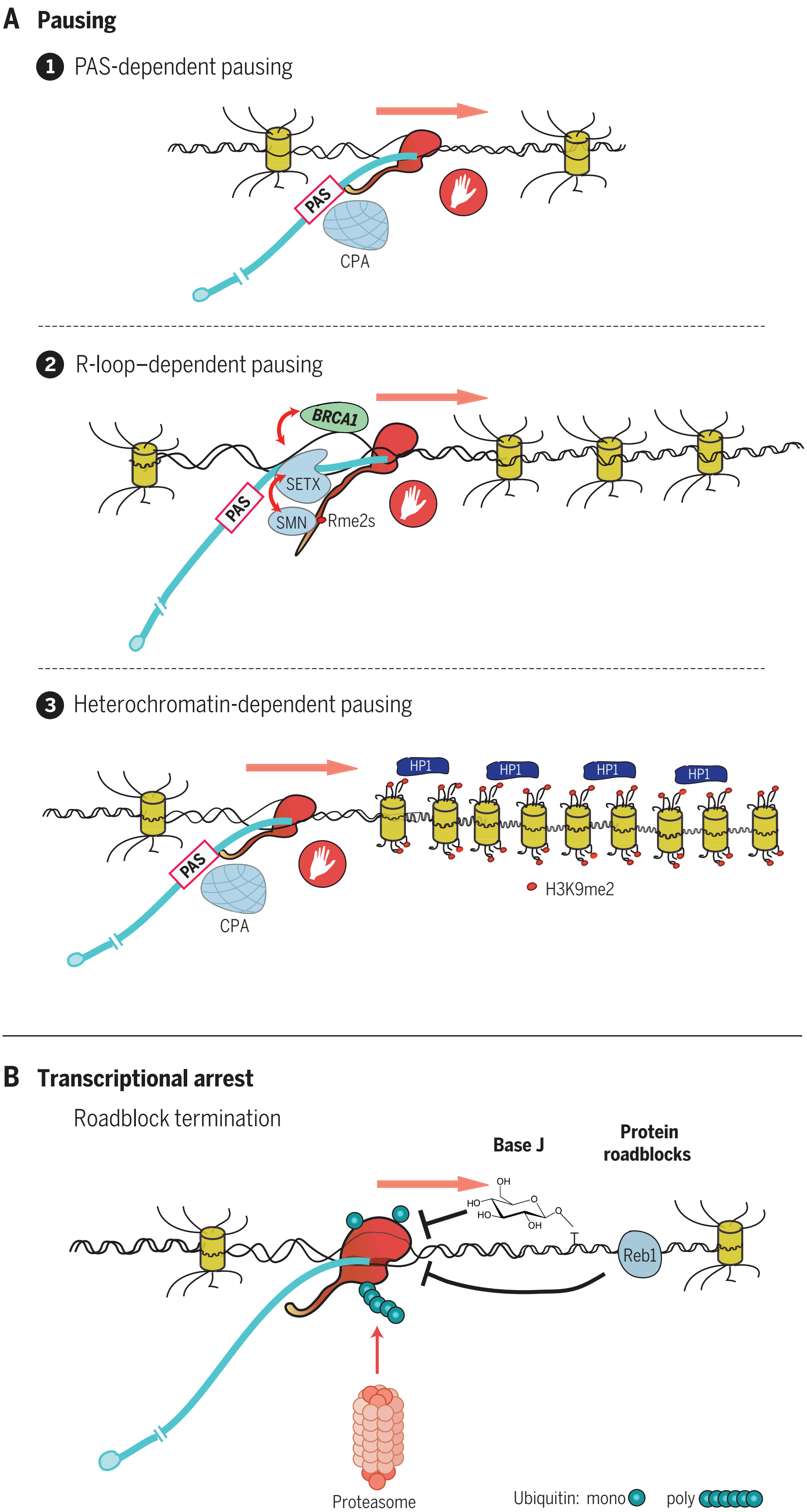


Fig. 1. Pausing or arresting Pol II. (**A**) Three different types of Pol II pausing induced by CPA recognition of the PAS, R-loop formation, and heterochromatin patches. Elongating Pol II (red) is shown transcribing the DNA template, with extruded, capped RNA transcript (blue) indicated. Nucleosomes are depicted by yellow barrels, with histone N-terminal tails indicated. Pol II CTD is shown as an extended tail. Red dots on the CTD and histone tails denote methylation. The hand denotes Pol II pausing. (**B**) Pol II arrested by base J or Reb1 DNA binding protein. Pol II is then ubiquitinated and degraded by the proteasome.

orientation. Transcript reads terminate precisely at J sites. Deletion of genes encoding the enzymes that perform this T-J conversion result in read-through transcription and ultimately cell death (33). By analogy to DNA damage in mammals, Pol II is likely released from base J roadblocks by ubiquitin-triggered proteolysis. It is, however, unclear how trypanosomes would tolerate such a high turnover of Pol II. Also, the selection mechanism that places base J at the end of TUs has yet to be determined (34).

In *Saccharomyces cerevisiae*, the DNA binding protein Reb1 is well known to act as a transcription factor for ribosomal protein-coding genes (35). Reb1 also binds promiscuously to intergenic sequences, where it restricts readthrough transcription, acting to prevent global transcriptional interference between genes. Like DNA damage and T-J–mediated termination, Reb1 termination causes Pol II arrest, again requiring ubiquitin-mediated Pol II destruction (36).

Torpedo termination

The 3′ end of eukaryotic mRNA is formed by an RNA processing mechanism whereby the CPA complex assembles onto the pre-mRNA PAS as it is extruded from the RNA exit channel of Pol II. This is facilitated by prior recruitment of CPA to the nearby Pol II CTD (37). Two models are widely cited for how PAS recognition triggers termination. One, dubbed the allosteric model (indicative of Pol II conformational change), proposes that elongating Pol II somehow senses its passage through a functional PAS, as described above (20). Likely, this is caused by the association of the very large CPA complex with Pol II CTD. This in turn induces a conformational change within the Pol II active site, resulting in first pausing and then Pol II release.

The alternative model (38, 39) relates to the nascent transcript still being synthesized by Pol II after cleavage at the PAS. The nuclear 5′-3′ exonuclease Xrn2 is recruited to PAS and progressively degrades this downstream transcript in kinetic competition with ongoing Pol II elongation. When Xrn2 catches up with Pol II, this acts as a molecular trigger to release Pol II from the DNA template. Clearly, pausing of Pol II—caused, for instance, by R-loop–mediated heterochromatic marks (Fig. 1A)—will enhance Xrn2-mediated termination (29). This proposed mechanism is evocatively named the torpedo model where, in naval vernacular, Pol II is the battleship and Xrn2 the torpedo. Direct evidence came from depletion of Rat1 in *S. cerevisiae* or Xrn2 in mammalian cells provoking a substantial loss in termination (40, 41). Rat1 interacts with Rai1, which possesses both pyrophosphatase and some 5′-3′ exonuclease activity (42), so that both together promote more efficient RNA degradation. A third member of this torpedo complex, called Rtt103, possesses a CTD interaction domain (CID), possibly accounting for Rat1 recruitment to Pol II (40). Other CPA factors may also aid Rat1 recruitment, including Pcf11, which also possesses a CID (43). In general, the depletion of Xrn2 by RNAi technology in mammalian cells produced varying degrees of termination defect, leading to the view that other factors might cooperate with Xrn2 or Rat1 to achieve more efficient termination (15). One such factor is the RNA:DNA helicase Sen1 in yeast or Senataxin in mammals, which may act to expose the cleaved downstream RNA product to Xrn2 degradation. This may be particularly important if RNA is entwined with the DNA template in an R-loop structure (RNA:DNA hybrid) (25, 26). Interestingly, Sen1 displays termination activity independently of RNA degradation (44), indicating that the simple act of unraveling RNA structure is enough to destabilize Pol II and promote termination. Such a mechanism is very similar to the termination activity of *E. coli* Rho. Although depletion of Xrn2 by RNAi often causes only a marginal termination defect, combining RNAi treatment with expression of dominant negative Xrn2 (active-site mutant) gives a satisfyingly large termination defect for most protein-coding genes (45). It appears that depletion of Xrn2 cellular levels (by RNAi) may not adequately reduce levels of Xrn2 actively engaged in termination.

An underlying feature of productive termination is that 3′-end cleavage of the nascent transcript at the PAS facilitates termination by allowing Xrn2 "torpedo" action. The *S. cerevisiae* RNAse III Rnt1 recognizes specific hairpin structures and carries out a double endonuclease cut across the hairpin. Several yeast genes use Rnt1 cleavage as alternative 3′-end processing events that allow PAS-independent Xrn2-mediated termination (46, 47). A similar mechanism operates on lncRNA-derived primary microRNAs (lnc-pri-miRNAs), which are independently transcribed rather than more usually residing in the introns of protein-coding genes. lnc-pri-miRNAs are cotranscriptionally cleaved by Drosha, a distant relative of Rnt1. RNA cleavage releases pre-miRNAs, which are then exported from the nucleus and converted into miRNAs by another Rnt1 relative called Dicer. Notably, Drosha cleavage not only generates pre-miRNA but also promotes lncRNA gene termination further downstream, again likely involving the action of Xrn2 degradation (13, 48). To underline the effectiveness of Rnt1 cleavage as a termination mechanism, Pol I transcription is also terminated by Rnt1 cleavage of a hairpin at the 3′ end of rRNA genes. This again elicits termination by a combination of Xrn2 and Sen1 action (49).

A further category of termination has been uncovered wherein Pol II continues to generate an extended transcript multiple kilobases into the gene 3′-flanking region after passage of the PAS. Termination eventually occurs, coincident with a terminal cotranscriptionally cleaved transcript (CoTC sequence), which generates cleavage products closely associated with Pol II. These are degraded by Xrn2 and so promote termination of Pol II in their vicinity. CoTC termination still requires the presence of an upstream PAS. This suggests that the conformational change induced by CPA recognition of the PAS is required for downstream CoTC-mediated termination (50). Cleavage at the PAS to finally release polyadenylated mRNA may occur following CoTC-mediated termination, effectively after the Pol II complex, with its still associated pre-mRNA, is released into the nucleoplasm (51).

Transcriptional backtracking to promote termination

All RNA polymerases can transcribe in both forward and reverse directions on the DNA template. Forward movement results in template-dependent RNA synthesis, with the nascent transcript emerging from the RNA exit channel. Backward movement (backtracking) results in extrusion of the already synthesized nascent transcript out of the secondary channel (also referred to as the nucleotide entry channel) (52). During transcriptional elongation, such backtracking is widely used by Pol II as a proofreading mechanism. The general transcription factor TFIIS enhances an intrinsic endonuclease activity of Pol II that promotes cleavage of the mismatched extruded RNA. This allows transcription to resume, reinstating the correct nucleotide into the nascent RNA (53, 54). For both *E. coli* RNA polymerase and eukaryotic Pol III, backtracking has been directly implicated in termination. For intrinsic termination in *E. coli*, a model is envisaged wherein oligo(U) sequences promote polymerase pausing, which then favors backtracking. If an RNA hairpin forms on the upstream transcript, it can be forced into the RNA exit channel, which in turn triggers a conformational change in the polymerase that promotes its release from the DNA template (2). Pol III appears to adopt a similar strategy, as transcript 3′ ends are normally oligo(U) sequences, which pause the polymerase and so encourage backtracking. If backward polymerase movement encounters a hairpin structure, then termination ensues (55) (Fig. 2A).

Pol II may employ a related backtracking mechanism to promote termination. In the yeast *Schizosaccharomyces pombe*, inactivation of the RNA exosome displays a clear general termination defect, which is counteracted by simultaneous loss of TFIIS (56). Because the multisubunit exosome possesses two separate 3′-5′exonucleases, one of these may act on the extruded RNA formed by backtracking. This may push the polymerase further backward and by so doing, induce conformational changes in the Pol II active site that promote termination. Further evidence for such a mechanism comes from in vitro termination experiments using purified yeast Pol II, Rat1, and Rai1 together with immobilized DNA templates, where transcription is artificially blocked by omitting specific nucleotides (57). Although earlier in vitro experiments failed to observe Pol II termination with these minimal components (58), substantial termination is observed if the wrong nucleotide is forced onto the transcript 3′ end, so arresting Pol II at this mismatch position. This has the effect of inducing Pol II backtracking and also, remarkably, promotes termination when coupled with degradation of the upstream transcript by Xrn2 up to the arrested Pol II (57). In effect, these experiments argue that the act of removing RNA up to or into the Pol II active site by either degradation of backtracked extruded transcript (reverse torpedo) or degradation of upstream RNA up to backtracked Pol II (forward

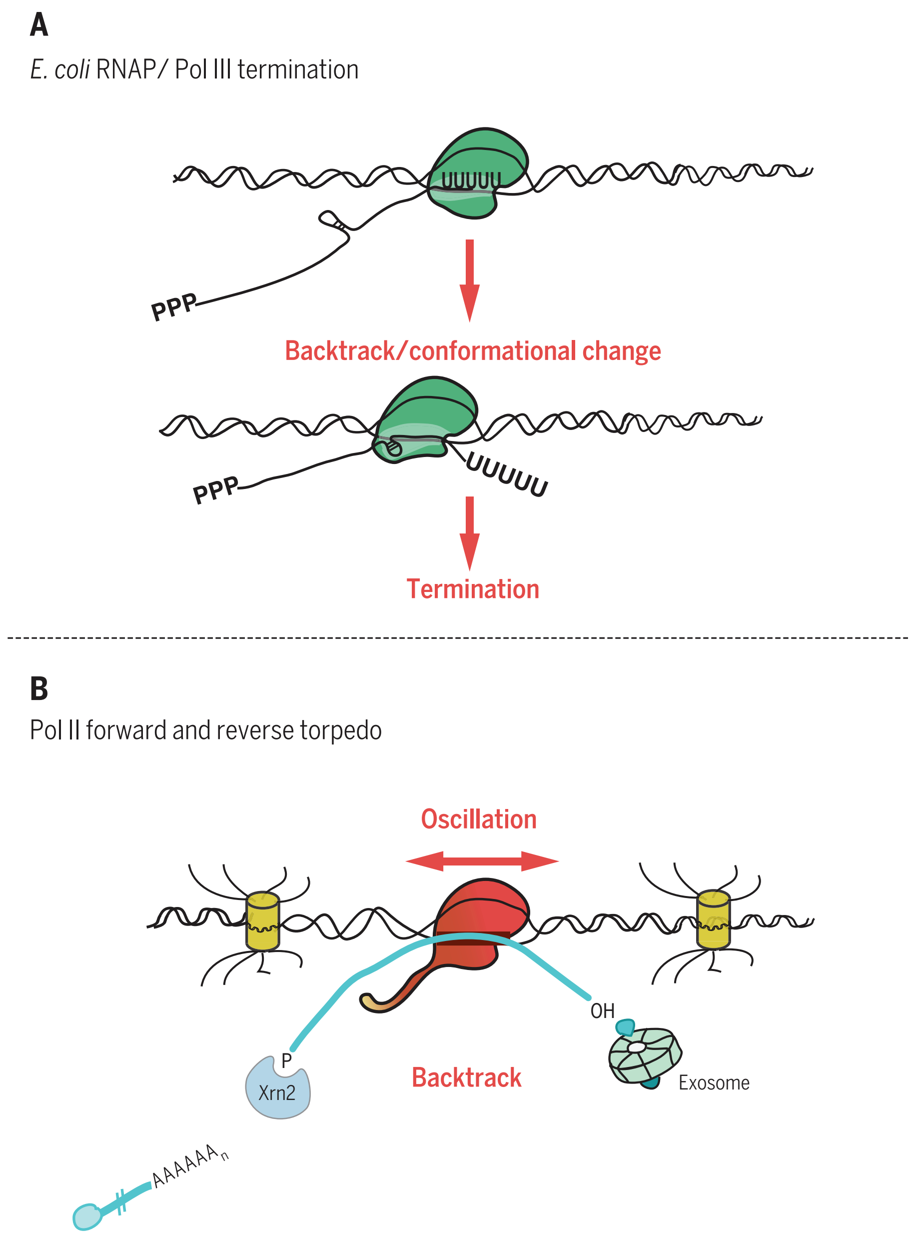


Fig. 2. Pol II backtracking. (**A**) Bacterial RNA polymerase or Pol III terminates at an oligo(U) transcript, which pauses polymerase and promotes backtracking. An upstream RNA hairpin is forced into polymerase active site, inducing a conformational change that results in termination. (**B**) Pol II moves forward to synthesize or backward to extrude transcript (oscillation). Forward transcript, once cleaved at PAS to release mRNA, is then degraded by Xrn2. Backtracked transcript is degraded by the exosome. Removal of RNA up to Pol II (forward or reverse torpedo) induces termination.

torpedo) induces conformational changes to Pol II that promote termination (Fig. 2B).

Premature termination versus transcriptional elongation

Saccharomyces cerevisiae possesses a secondary Pol II termination mechanism that operates on short Pol II transcripts. This involves the NRD complex (Nrd1, Nab3, and Sen1) (*59*), which promotes termination of ncRNA, particularly that derived from antisense promoter activity associated with promoters of protein-coding genes. It also functions on genes encoding small stable RNA, snRNA, and snoRNA and plays a major role in regulating a subset of protein-coding genes by promoting their premature termination. NRD promotes termination in a sequence-specific manner (through RNA recognition domains on Nrd1 and Nab3) and recruits the exosome to rapidly degrade these transcripts. This degradation will occur unless the transcripts are protected by RNA binding proteins that package snRNA and snoRNA into functional splicing or RNA modification complexes, respectively.

Although mammalian genes have no clear counterpart to NRD, most protein-coding genes display substantial promotor proximal pausing. This is manifested by an accumulation of actively transcribing Pol II localized to the first few hundred nucleotides of the gene (*60*). In contrast, *S. cerevisiae* genes show little Pol II pausing at the promoter. It appears that the mammalian transcriptome, possibly because of its greater complexity, has evolved mechanisms of transcriptional regulation more focused on postinitiation events. However, if these early transcripts are not intermediates waiting to be converted into full-length gene transcripts and are therefore abortive, they need to be terminated. Bona fide termination clearly operates on these TSS (transcription start site) transcripts. First, decapping of the transcript can occur by Dcp1 action, followed by Xrn2 degradation (*61*). Second, mis-spliced transcripts are somehow detected by nuclear surveillance and are similarly degraded by decapping and Xrn2 degradation (*62*). Promoter-proximal PASs are also thought to be actively recognized by CPA and Xrn2 to promote early termination. Thus, depletion of either CPA components or Xrn2 increases levels of TSS-associated transcripts (*15*). Some of these early terminated TSS transcripts form hairpin structures, which directly act as functional pre-microRNA without the involvement of the microprocessor (*63*). It is apparent that TSS-associated termination may be regulated by 5′ splice sites, which block promoter proximal PASs and thereby favor continued elongation into the gene body. This phenomenon was first shown in viruses such as HIV-1 (*64*), but has also been revealed as a general mechanism that acts to block cryptic PAS recognition and consequent premature termination. In particular, depletion of U1 snRNA activates TSS-proximal PASs as well as numerous PASs present across genes, often within their extensive intronic regions (*65*).

Transcription elongation is tightly controlled across genes (*66*). Specific check points operate to enforce premature termination that acts to prevent inappropriate transcription (Fig. 3). Early transcription elongation is restricted by two negative elongation factors, DRB-sensitivity–inducing factor (DSIF) and negative elongation factor (NELF), which are regulated by the major Pol II CTD Ser^2 kinase Cdk9 (cyclin-dependent kinase 9). As well as acting on the CTD, this kinase further phosphorylates DSIF and NELF. Thus, when Cdk9 is experimentally inhibited by drugs such as 5,6-dichloro-1-B-D-ribofuranosylbenzimidazole (DRB) or KM05382 (KM), a substantial increase in TSS-proximal transcripts is observed, with greatly reduced transcription downstream into the gene body (*67*). Once Pol II escapes from TSS-proximal checkpoints and elongation is fully under way, then numerous elongation factors come into play. These promote efficient transcription across the TU, be it a modest 1-kb or much longer 1-Mb gene. Many of these elongation factors act to remodel nucleosomes encountered by elongating Pol II, as well as to coordinate efficient and often regulated (alternative) intron splicing (*66*). Recently, a further example of regulated premature termination

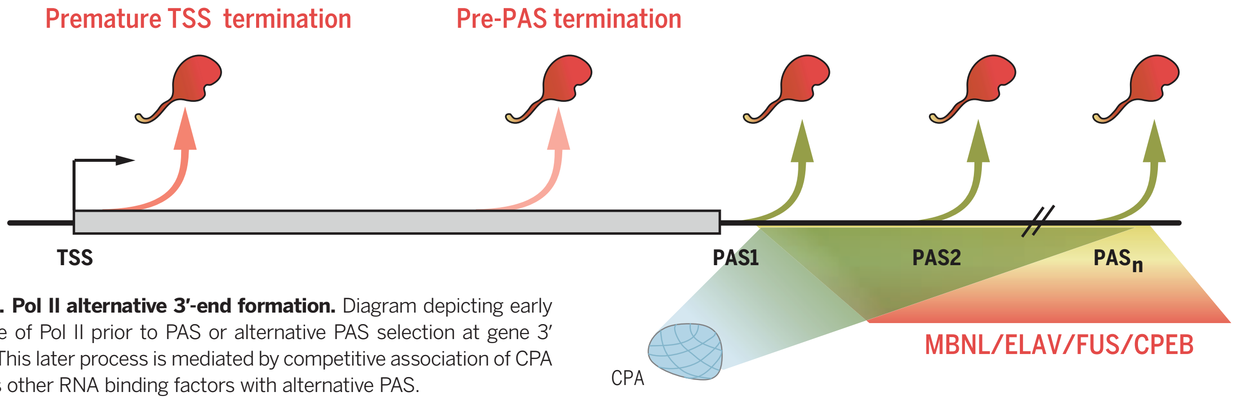


Fig. 3. Pol II alternative 3′-end formation. Diagram depicting early release of Pol II prior to PAS or alternative PAS selection at gene 3′ ends. This later process is mediated by competitive association of CPA versus other RNA binding factors with alternative PAS.

has been observed near the end of gene TUs (3′-end checkpoint) that may act as a final control mechanism to prevent the production of a translatable but potentially flawed mRNA. Like the TSS checkpoint, the 3′-end checkpoint is controlled by Cdk9 activity. In this case, the substrate may be Xrn2, which is substantially activated by Cdk9-mediated phosphorylation (68). Inhibition of Cdk9 causes a nonproductive termination mechanism that still appears to use component parts of CPA, but apparently does not allow the formation of functional polyadenylated mRNA (67).

Alternative polyadenylation and termination

Termination at gene 3′ ends is also subject to intense regulation. Many mRNAs possess variable lengths of 3′-untranslated sequence defined by the selective usage of different PASs (69). Because mRNA 3′UTRs define mRNA cytoplasmic functions, including RNA stability, translatability, and localization, the use of alternative poly(A) sites (APA) can constitute a key regulatory process in gene expression. However, the differential stability of different mRNA 3′UTR isoforms in the cytoplasm must be distinguished from actual PAS selection during pre-mRNA synthesis. Analyzing total cellular mRNA isoform levels will mainly show up differential stability, whereas analysis of nuclear mRNA isoform levels more closely reflects PAS selection. Indeed, nuclear APA shows rather less variation and consequently gene regulation than initially envisaged (70). An issue often overlooked is whether APA constitutes alternative RNA processing alone or whether selective termination defines which PAS is used. Furthermore, APA may also occur at more internal positions within a gene where it will result in truncated mRNA, which in some cases can be translated into shorter protein isoforms with important biological functions. A classic example of this phenomenon is the alternative use of an internal PAS in immunoglobulin M (IgM) that generates secreted antibody versus a distal PAS that produces membrane-bound antibody. This APA switch is regulated by levels of CstF-64, which is depleted in early B cells and favors distal PAS selection (71). In most examples of APA (as in IgM), the first (proximal) PAS will have a weaker match to the consensus sequence features than the downstream (distal) PAS (72). As mentioned above, Pol II pausing will also have notable effects here, as pausing between proximal and distal PAS favors the proximal PAS. Once a PAS is selected, this will trigger Pol II termination downstream and so preclude distal PAS usage (72).

Several specific conditions are known to affect APA. The kinetics of transcription can have a major effect. Also, rapidly proliferating cells favor proximal PAS usage (69). How a gene promoter recruits elongation factors such as the PAF complex can further influence downstream Pol II elongation and consequent PAS selection (73). Finally, directly slowing down Pol II elongation by specific Pol II mutation can favor proximal PAS usage (74). Either slowing down or speeding up Pol II elongation by specific Pol II mutation can be seen to globally shorten or extend Pol II TUs, implying effects on both APA and coupled termination (45). The availability of CPA factors can also affect APA. Systematic depletion of individual CPA components by RNAi treatment can influence APA (75, 76). For example, depletion of CFI components and PABPN1 reduces, whereas Pcf11 and Fip1 depletion enhances, distal PAS usage. However, it is unclear whether modulation of CPA factor concentration occurs naturally. CF1m25 is overexpressed in glioblastoma cells, resulting in global proximal PAS selection (77). Also, PABN1 can be lost in a rare form of muscular dystrophy called OPMD (oculopharyngeal muscular dystrophy), where a triplet expansion in the gene generates an inactive protein containing N-terminal polyalanine. Loss of PABN1 causes a clear APA shift to distal PAS usage (78, 79). U1snRNA, already mentioned as a means to block premature PAS usage at upstream gene positions, can also affect APA. Thus. lowering U1snRNA levels favors proximal PAS usage. Possibly in activated neurons, U1snRNA may be sufficiently depleted to favor proximal PAS usage (80). A series of non-CPA RNA binding factors are known to be tightly regulated in amount and/or cellular location. These include CPEB, FUS, ELAV, and MBNL (81–84). All of these factors display some degree of RNA binding specificity, and where their binding sites are near PAS, may directly influence CPA binding and recruitment (Fig. 3).

The actual mechanism by which APA may be regulated in these diverse situations remains largely unknown. However, the consequences of an mRNA having a more or less extended 3′UTR clearly relates to key functions of these mRNA regulatory regions, including microRNA binding sites and RNA stability elements. Notably, the actual selection of different PAS can also result in the recruitment of protein factors to the 3′UTR, which, when the mRNA is translated, can directly associate with the nascent protein to modulate its function (85).

Misregulated termination

Transcriptional termination is a highly robust process capable of stopping the Pol II juggernaut at gene 3′ ends. It is therefore paradoxical that this basic mechanism appears to be readily subverted for particular cellular or viral needs. First, a notable feature of repressed termination can be seen in the primary Piwi-interacting RNA (piRNA) clusters of *Drosophila*. These clusters harbor a wide range of transposon-related sequences and in *Drosophila* are often transcribed on both DNA strands. Primary piRNAs are processed into small, 20-nucleotide piRNAs by dedicated RNA processing enzymes, including specific argonaute proteins. piRNAs act to block the spread of transposons and retroposons in a wide range of eukaryotes by targeting transposon-derived mRNA, using a mechanism analogous to that of miRNA (86). Notably, piRNA clusters lack independent promoters, but are positioned between convergent genes so that they are transcribed by readthrough transcription (87–89). Owing to dsRNA formation in these clusters, the primary piRNA chromatin is marked by H3K9me3, normally a feature of repressed, heterochromatin protein 1 (HP1)–associated heterochromatin. However, these TUs are bound by a distinct HP1 protein called Rhino, which is associated with other proteins including a Rai1 homolog called Cutoff. Rhino enhances rather than represses Pol II transcription across the piRNA TUs, and Cutoff further aids this readthrough transcription process by blocking Pol II termination. Cutoff may act as a dominant negative regulator of Xrn2 (45). Not only are the normal termination sites of the flanking convergent genes blocked, thereby promoting efficient Pol II readthrough transcription across both strands of the primary piRNA TUs, but terminators prevalent in transposon termini are similarly restricted (Fig. 4A).

Further examples of blocked termination come from cells responding to stress. The artificial

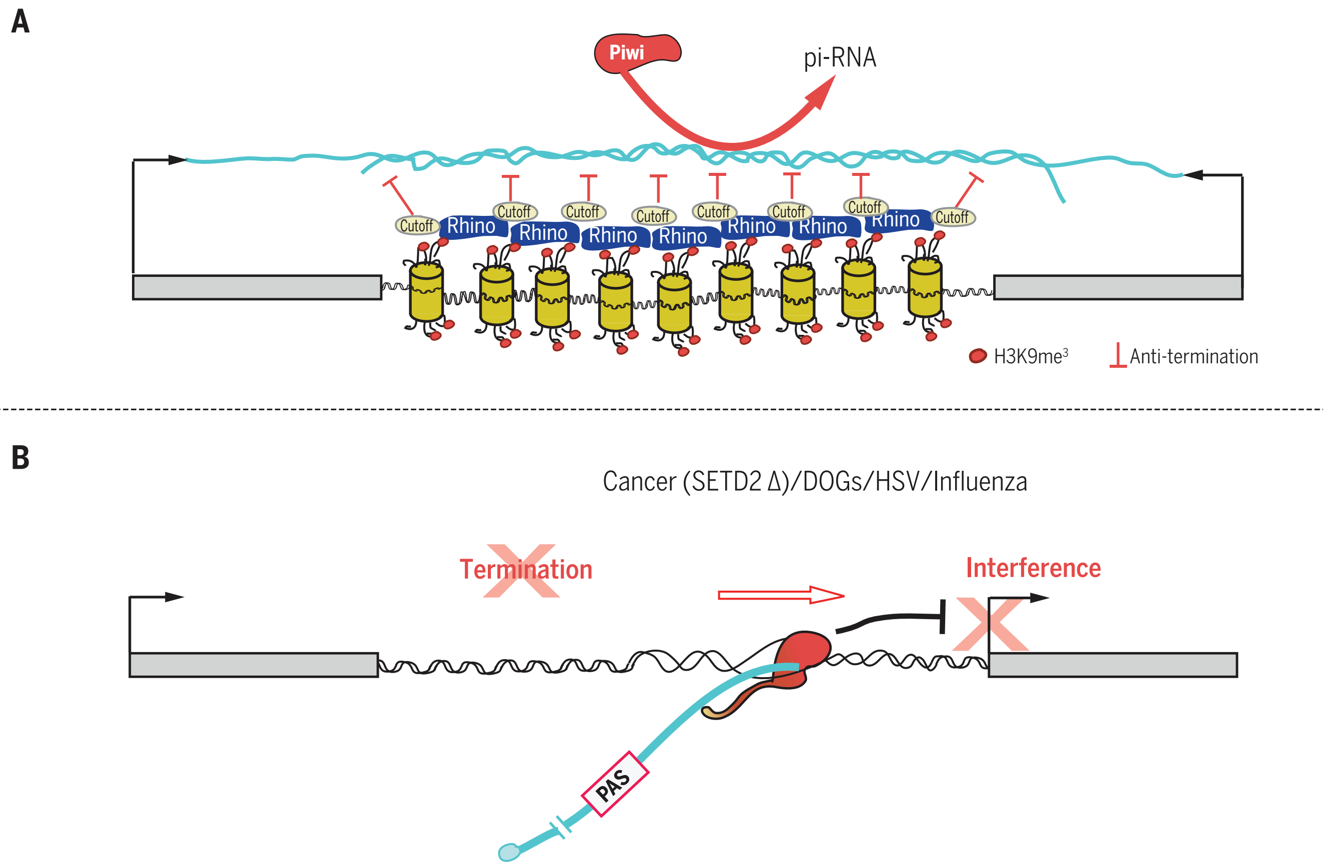


Fig. 4. Regulated or misregulated Pol II termination. (**A**) Primary piRNA clusters in *Drosophila* are often positioned between convergent genes. dsRNA induces heterochromatic histone tail modification (H3K9me3). This in turn recruits the HP1-like factor Rhino together with Cutoff, an anti-terminator that promotes readthrough transcription. (**B**) Inactivation of termination by cancer mutation (SETD2 mutation), osmotic stress, or viral infection all induce Pol II readthrough and interference with downstream gene expression.

induction of osmotic stress in cultured cells results in a large number of genes failing to terminate efficiently at the normal gene 3′ end. Instead, readthrough transcripts are detectable that may extend through intergenic regions and invade downstream genes. These "downstream of gene" transcripts (DOGs), though potentially deleterious owing to interference effects, have been postulated to possess protective features for the overall integrity of the nucleus under cellular stress (*90*). Similarly, cancer cells display complex readthrough transcription profiles that may be related to DOGs (*91*–*93*). In renal cancer, mutations in the methyltransferase gene *SETD2* correlate with readthrough transcription profiles (*94*). Setd2 adds the histone H3K36me3 mark to genic nucleosomes, which is required for Pol II elongation and termination. Setd2 also cooperates with Pol II elongation factors and facilitates Pol II CTD phospho-Ser^2 formation, which will ultimately lead to CPA recruitment and termination (*95*).

Viral infection may be considered an extreme form of cellular stress. Remarkably, at least two viruses are known to drastically perturb the termination efficiency of their host genomes. Thus, influenza virus essentially blocks host transcription termination, genome-wide. In detail, the viral protein NS1 has high affinity for CPSF-30 and through this interaction destroys CPA complex integrity. This leads to a general loss in 3′-end processing with commensurate readthrough transcription (*96*). Herpes simplex virus, upon infecting host cells, similarly causes a massive misregulation of host gene transcription. Extensive readthrough transcription occurs with all the associated interference and mis-splicing of these extended transcripts (*97*) (Fig. 4B). The destruction of normal regulated Pol II termination in host cells infected with these common human viruses likely explains their pathogenicity.

Conclusions

This Review charts our increasing understanding of how transcriptional termination affects many aspects of eukaryotic gene expression. Far from acting as a constitutive mechanism to separate TUs across the genome, termination can be seen as an intricate process that displays remarkable flexibility and regulatory potential. At the beginning of the gene, termination regulates transcript release into productive elongation. It also acts as a checkpoint to prevent the synthesis of defective mRNA, which could be translated into a toxic (dominant negative) protein. At the end of the gene, termination dictates which mRNA isoform is formed by APA, thereby conferring selective expression properties on the mRNA. Finally, termination can be overridden to adjust cells to stress conditions or to adapt cells into a more pliable host for viral replication. It is likely that future analysis of the termination process has yet more surprises in store.

REFERENCES AND NOTES

1. T. Platt, Transcription termination and the regulation of gene expression. *Annu. Rev. Biochem.* **55**, 339–372 (1986). doi: 10.1146/annurev.bi.55.070186.002011; pmid: 3527045
2. V. Epshtein, C. J. Cardinale, A. E. Ruckenstein, S. Borukhov, E. Nudler, An allosteric path to transcription termination. *Mol. Cell* **28**, 991–1001 (2007). doi: 10.1016/j.molcel.2007.10.011; pmid: 18158897
3. J. P. Richardson, Preventing the synthesis of unused transcripts by Rho factor. *Cell* **64**, 1047–1049 (1991). doi: 10.1016/0092-8674(91)90257-Y; pmid: 2004415
4. V. Epshtein, D. Dutta, J. Wade, E. Nudler, An allosteric mechanism of Rho-dependent transcription termination. *Nature* **463**, 245–249 (2010). doi: 10.1038/nature08669; pmid: 20075920
5. T. M. Henkin, C. Yanofsky, Regulation by transcription attenuation in bacteria: How RNA provides instructions for transcription termination/antitermination decisions. *BioEssays* **24**, 700–707 (2002). doi: 10.1002/bies.10125; pmid: 12210530
6. B. S. Gloss, M. E. Dinger, The specificity of long noncoding RNA expression. *Biochim. Biophys. Acta* **1859**, 16–22 (2015). doi: 10.1016/j.bbagrm.2015.08.005; pmid: 26297315
7. I. H. Greger, N. J. Proudfoot, Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of Saccharomyces cerevisiae. *EMBO J.*

17, 4771–4779 (1998). doi: 10.1093/emboj/17.16.4771; pmid: 9707436
8. K. E. Shearwin, B. P. Callen, J. B. Egan, Transcriptional interference—a crash course. Trends Genet. 21, 339–345 (2005). doi: 10.1016/j.tig.2005.04.009; pmid: 15922833
9. M. Gullerova, N. J. Proudfoot, Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Cell 132, 983–995 (2008). doi: 10.1016/j.cell.2008.02.040; pmid: 18358811
10. E. M. Prescott, N. J. Proudfoot, Transcriptional collision between convergent genes in budding yeast. Proc. Natl. Acad. Sci. U.S.A. 99, 8796–8801 (2002). doi: 10.1073/pnas.132270899; pmid: 12077310
11. D. J. Hobson, W. Wei, L. M. Steinmetz, J. Q. Svejstrup, RNA polymerase II collision interrupts convergent transcription. Mol. Cell 48, 365–374 (2012). doi: 10.1016/j.molcel.2012.08.027; pmid: 23041286
12. H. Gaillard, T. García-Muse, A. Aguilera, Replication stress and cancer. Nat. Rev. Cancer 15, 276–289 (2015). doi: 10.1038/nrc3916; pmid: 25907220
13. A. Dhir, S. Dhir, N. J. Proudfoot, C. L. Jopling, Microprocessor mediates transcriptional termination of long noncoding RNA transcripts hosting microRNAs. Nat. Struct. Mol. Biol. 22, 319–327 (2015). doi: 10.1038/nsmb.2982; pmid: 25730776
14. L. J. Core, J. J. Waterfall, J. T. Lis, Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322, 1845–1848 (2008). doi: 10.1126/science.1162228; pmid: 19056941
15. T. Nojima et al., Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. Cell 161, 526–540 (2015). doi: 10.1016/j.cell.2015.03.027; pmid: 25910207
16. A. Vannini, P. Cramer, Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Mol. Cell 45, 439–446 (2012). doi: 10.1016/j.molcel.2012.01.023; pmid: 22365827
17. S. McCracken et al., The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature 385, 357–361 (1997). doi: 10.1038/385357a0; pmid: 9002523
18. M. Heidemann, C. Hintermair, K. Voß, D. Eick, Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. Biochim. Biophys. Acta 1829, 55–62 (2013). doi: 10.1016/j.bbagrm.2012.08.013; pmid: 22982363
19. J. P. Hsin, J. L. Manley, The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. 26, 2119–2137 (2012). doi: 10.1101/gad.200303.112; pmid: 23028141
20. H. Zhang, F. Rigo, H. G. Martinson, Poly(A) signal-dependent transcription termination occurs through a conformational change mechanism that does not require cleavage at the poly(A) site. Mol. Cell 59, 437–448 (2015). doi: 10.1016/j.molcel.2015.06.008; pmid: 26166703
21. A. Aguilera, T. García-Muse, R loops: From transcription byproducts to threats to genome stability. Mol. Cell 46, 115–124 (2012). doi: 10.1016/j.molcel.2012.04.009; pmid: 22541554
22. K. Skourti-Stathaki, N. J. Proudfoot, A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. 28, 1384–1396 (2014). doi: 10.1101/gad.242990.114; pmid: 24990962
23. S. Jimeno, A. G. Rondón, R. Luna, A. Aguilera, The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. EMBO J. 21, 3526–3535 (2002). doi: 10.1093/emboj/cdf335; pmid: 12093753
24. X. Li, J. L. Manley, Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell 122, 365–378 (2005). doi: 10.1016/j.cell.2005.06.008; pmid: 16096057
25. H. E. Mischo et al., Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol. Cell 41, 21–32 (2011). doi: 10.1016/j.molcel.2010.12.007; pmid: 21211720
26. K. Skourti-Stathaki, N. J. Proudfoot, N. Gromak, Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell 42, 794–805 (2011). doi: 10.1016/j.molcel.2011.04.026; pmid: 21700224
27. D. Y. Zhao et al., SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. Nature 529, 48–53 (2016). doi: 10.1038/nature16469; pmid: 26700805
28. E. Hatchi et al., BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol. Cell 57, 636–647 (2015). doi: 10.1016/j.molcel.2015.01.011; pmid: 25699710
29. K. Skourti-Stathaki, K. Kamieniarz-Gdula, N. J. Proudfoot, R-loops induce repressive chromatin marks over mammalian gene terminators. Nature 516, 436–439 (2014). doi: 10.1038/nature13787; pmid: 25296254
30. J. Q. Svejstrup, The interface between transcription and mechanisms maintaining genome integrity. Trends Biochem. Sci. 35, 333–338 (2010). doi: 10.1016/j.tibs.2010.02.001; pmid: 20194025
31. J. M. Santos-Pereira, A. Aguilera, R loops: New modulators of genome dynamics and function. Nat. Rev. Genet. 16, 583–597 (2015). doi: 10.1038/nrg3961; pmid: 26370899
32. A. Helmrich, M. Ballarino, L. Tora, Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Mol. Cell 44, 966–977 (2011). doi: 10.1016/j.molcel.2011.10.013; pmid: 22195969
33. H. G. van Luenen et al., Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in Leishmania. Cell 150, 909–921 (2012). doi: 10.1016/j.cell.2012.07.030; pmid: 22939620
34. P. A. Genest et al., Defining the sequence requirements for the positioning of base J in DNA using SMRT sequencing. Nucleic Acids Res. 43, 2102–2115 (2015). doi: 10.1093/nar/gkv095; pmid: 25662217
35. P. D. Hartley, H. D. Madhani, Mechanisms that specify promoter nucleosome location and identity. Cell 137, 445–458 (2009). doi: 10.1016/j.cell.2009.02.043; pmid: 19410542
36. J. Colin et al., Roadblock termination by reb1p restricts cryptic and readthrough transcription. Mol. Cell 56, 667–680 (2014). doi: 10.1016/j.molcel.2014.10.026; pmid: 25479637
37. N. Proudfoot, New perspectives on connecting messenger RNA 3′ end formation to transcription. Curr. Opin. Cell Biol. 16, 272–278 (2004). doi: 10.1016/j.ceb.2004.03.007; pmid: 15145351
38. S. Connelly, J. L. Manley, A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. 2, 440–452 (1988). doi: 10.1101/gad.2.4.440; pmid: 2836265
39. N. J. Proudfoot, How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem. Sci. 14, 105–110 (1989). doi: 10.1016/0968-0004(89)90132-1; pmid: 2658217
40. M. Kim et al., The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature 432, 517–522 (2004). doi: 10.1038/nature03041; pmid: 15565157
41. S. West, N. Gromak, N. J. Proudfoot, Human 5′ —> 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature 432, 522–525 (2004). doi: 10.1038/nature03035; pmid: 15565158
42. S. Xiang et al., Structure and function of the 5′—>3′ exoribonuclease Rat1 and its activating partner Rai1. Nature 458, 784–788 (2009). doi: 10.1038/nature07731; pmid: 19194460
43. D. D. Licatalosi et al., Functional interaction of yeast pre-mRNA 3′ end processing factors with RNA polymerase II. Mol. Cell 9, 1101–1111 (2002). doi: 10.1016/S1097-2765(02)00518-X; pmid: 12049745
44. O. Porrua, D. Libri, A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. Nat. Struct. Mol. Biol. 20, 884–891 (2013). doi: 10.1038/nsmb.2592; pmid: 23748379
45. N. Fong et al., Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. Mol. Cell 60, 256–267 (2015). doi: 10.1016/j.molcel.2015.09.026; pmid: 26474067
46. A. G. Rondón, H. E. Mischo, J. Kawauchi, N. J. Proudfoot, Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. Mol. Cell 36, 88–98 (2009). doi: 10.1016/j.molcel.2009.07.028; pmid: 19818712
47. G. Ghazal et al., Yeast RNase III triggers polyadenylation-independent transcription termination. Mol. Cell 36, 99–109 (2009). doi: 10.1016/j.molcel.2009.07.029; pmid: 19818713
48. M. Ballarino et al., Coupled RNA processing and transcription of intergenic primary microRNAs. Mol. Cell. Biol. 29, 5632–5638 (2009). doi: 10.1128/MCB.00664-09; pmid: 19667074
49. J. Kawauchi, H. Mischo, P. Braglia, A. Rondon, N. J. Proudfoot, Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. Genes Dev. 22, 1082–1092 (2008). doi: 10.1101/gad.463408; pmid: 18413718
50. S. West, N. J. Proudfoot, M. J. Dye, Molecular dissection of mammalian RNA polymerase II transcriptional termination. Mol. Cell 29, 600–610 (2008). doi: 10.1016/j.molcel.2007.12.019; pmid: 18342606
51. T. Nojima, M. Dienstbier, S. Murphy, N. J. Proudfoot, M. J. Dye, Definition of RNA polymerase II CoTC terminator elements in the human genome. Cell Reports 3, 1080–1092 (2013). doi: 10.1016/j.celrep.2013.03.012; pmid: 23562152
52. E. Nudler, RNA polymerase backtracking in gene regulation and genome instability. Cell 149, 1438–1445 (2012). doi: 10.1016/j.cell.2012.06.003; pmid: 22726433
53. S. Sigurdsson, A. B. Dirac-Svejstrup, J. Q. Svejstrup, Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. Mol. Cell 38, 202–210 (2010). doi: 10.1016/j.molcel.2010.02.026; pmid: 20417599
54. M. G. Izban, D. S. Luse, The RNA polymerase II ternary complex cleaves the nascent transcript in a 3′——5′ direction in the presence of elongation factor SII. Genes Dev. 6, 1342–1356 (1992). doi: 10.1101/gad.6.7.1342; pmid: 1378419
55. S. Nielsen, Y. Yuzenkova, N. Zenkin, Mechanism of eukaryotic RNA polymerase III transcription termination. Science 340, 1577–1580 (2013). doi: 10.1126/science.1237934; pmid: 23812715
56. J. F. Lemay et al., The RNA exosome promotes transcription termination of backtracked RNA polymerase II. Nat. Struct. Mol. Biol. 21, 919–926 (2014). doi: 10.1038/nsmb.2893; pmid: 25240800
57. J. Park, M. Kang, M. Kim, Unraveling the mechanistic features of RNA polymerase II termination by the 5′-3′ exoribonuclease Rat1. Nucleic Acids Res. 43, 2625–2637 (2015). doi: 10.1093/nar/gkv133; pmid: 25722373
58. S. Dengl, P. Cramer, Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II in vitro. J. Biol. Chem. 284, 21270–21279 (2009). doi: 10.1074/jbc.M109.013847; pmid: 19535338
59. H. E. Mischo, N. J. Proudfoot, Disengaging polymerase: Terminating RNA polymerase II transcription in budding yeast. Biochim. Biophys. Acta 1829, 174–185 (2013). doi: 10.1016/j.bbagrm.2012.10.003; pmid: 23085255
60. K. Adelman, J. T. Lis, Promoter-proximal pausing of RNA polymerase II: Emerging roles in metazoans. Nat. Rev. Genet. 13, 720–731 (2012). doi: 10.1038/nrg3293; pmid: 22986266
61. K. Brannan et al., mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol. Cell 46, 311–324 (2012). doi: 10.1016/j.molcel.2012.03.006; pmid: 22483619
62. L. Davidson, A. Kerr, S. West, Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. EMBO J. 31, 2566–2578 (2012). doi: 10.1038/emboj.2012.101; pmid: 22522706
63. M. Xie et al., Mammalian 5′-capped microRNA precursors that generate a single microRNA. Cell 155, 1568–1580 (2013). doi: 10.1016/j.cell.2013.11.027; pmid: 24360278
64. M. P. Ashe, P. Griffin, W. James, N. J. Proudfoot, Poly(A) site selection in the HIV-1 provirus: Inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. Genes Dev. 9, 3008–3025 (1995). doi: 10.1101/gad.9.23.3008; pmid: 7498796
65. D. Kaida et al., U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. Nature 468, 664–668 (2010). doi: 10.1038/nature09479; pmid: 20881964
66. H. Kwak, J. T. Lis, Control of transcriptional elongation. Annu. Rev. Genet. 47, 483–508 (2013). doi: 10.1146/annurev-genet-110711-155440; pmid: 24050178
67. C. Laitem et al., CDK9 inhibitors define elongation checkpoints at both ends of RNA polymerase II-transcribed genes. Nat. Struct. Mol. Biol. 22, 396–403 (2015). pmid: 25849141
68. M. Sansó et al., P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. Genes Dev. 30, 117–131 (2016). doi: 10.1101/gad.269589.115; pmid: 26728557
69. B. Tian, J. L. Manley, Alternative cleavage and polyadenylation: The long and short of it. Trends Biochem. Sci. 38, 312–320 (2013). doi: 10.1016/j.tibs.2013.03.005; pmid: 23632313
70. J. Neve et al., Subcellular RNA profiling links splicing and nuclear DICER1 to alternative cleavage and polyadenylation. Genome Res. 26, 24–35 (2016). doi: 10.1101/gr.193995.115; pmid: 26546131
71. Y. Takagaki, R. L. Seipelt, M. L. Peterson, J. L. Manley, The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell 87, 941–952 (1996). doi: 10.1016/S0092-8674(00)82000-0; pmid: 8945520
72. N. J. Proudfoot, Ending the message: Poly(A) signals then and now. Genes Dev. 25, 1770–1782 (2011). doi: 10.1101/gad.17268411; pmid: 21896654
73. T. Nagaike et al., Transcriptional activators enhance polyadenylation of mRNA precursors. Mol. Cell 41, 409–418 (2011). doi: 10.1016/j.molcel.2011.01.022; pmid: 21329879

74. P. A. Pinto *et al*., RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J.* **30**, 2431–2444 (2011). doi: 10.1038/emboj.2011.156; pmid: 21602789
75. W. Li *et al*., Systematic profiling of poly(A)+ transcripts modulated by core 3′ end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLOS Genet.* **11**, e1005166 (2015). doi: 10.1371/journal.pgen.1005166; pmid: 25906188
76. G. Martin, A. R. Gruber, W. Keller, M. Zavolan, Genome-wide analysis of pre-mRNA 3′ end processing reveals a decisive role of human cleavage factor I in the regulation of 3′ UTR length. *Cell Reports* **1**, 753–763 (2012). doi: 10.1016/j.celrep.2012.05.003; pmid: 22813749
77. C. P. Masamha *et al*., CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* **510**, 412–416 (2014). pmid: 24814343
78. M. Jenal *et al*., The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149**, 538–553 (2012). doi: 10.1016/j.cell.2012.03.022; pmid: 22502866
79. E. de Klerk *et al*., Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucleic Acids Res.* **40**, 9089–9101 (2012). doi: 10.1093/nar/gks655; pmid: 22772983
80. M. G. Berg *et al*., U1 snRNP determines mRNA length and regulates isoform expression. *Cell* **150**, 53–64 (2012). doi: 10.1016/j.cell.2012.05.029; pmid: 22770214
81. F. A. Bava *et al*., CPEB1 coordinates alternative 3′-UTR formation with translational regulation. *Nature* **495**, 121–125 (2013). doi: 10.1038/nature11901; pmid: 23434754
82. A. Masuda *et al*., Position-specific binding of FUS to nascent RNA regulates mRNA length. *Genes Dev.* **29**, 1045–1057 (2015). doi: 10.1101/gad.255737.114; pmid: 25995189
83. V. Hilgers, S. B. Lemke, M. Levine, ELAV mediates 3′ UTR extension in the Drosophila nervous system. *Genes Dev.* **26**, 2259–2264 (2012). doi: 10.1101/gad.199653.112; pmid: 23019123
84. R. Batra *et al*., Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell* **56**, 311–322 (2014). doi: 10.1016/j.molcel.2014.08.027; pmid: 25263597
85. B. D. Berkovits, C. Mayr, Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. *Nature* **522**, 363–367 (2015). doi: 10.1038/nature14321; pmid: 25896326
86. E. M. Weick, E. A. Miska, piRNAs: From biogenesis to function. *Development* **141**, 3458–3471 (2014). doi: 10.1242/dev.094037; pmid: 25183868
87. F. Mohn, G. Sienski, D. Handler, J. Brennecke, The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila. *Cell* **157**, 1364–1379 (2014). doi: 10.1016/j.cell.2014.04.031; pmid: 24906153
88. Z. Zhang *et al*., The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* **157**, 1353–1363 (2014). doi: 10.1016/j.cell.2014.04.030; pmid: 24906152
89. A. Le Thomas *et al*., Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes Dev.* **28**, 1667–1680 (2014). doi: 10.1101/gad.245514.114; pmid: 25085419
90. A. Vilborg, M. C. Passarelli, T. A. Yario, K. T. Tycowski, J. A. Steitz, Widespread Inducible Transcription Downstream of Human Genes. *Mol. Cell* **59**, 449–461 (2015). doi: 10.1016/j.molcel.2015.06.016; pmid: 26190259
91. C. A. Maher *et al*., Transcriptome sequencing to detect gene fusions in cancer. *Nature* **458**, 97–101 (2009). doi: 10.1038/nature07638; pmid: 19136943
92. K. Kannan *et al*., Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9172–9177 (2011). doi: 10.1073/pnas.1100489108; pmid: 21571633
93. K. E. Varley *et al*., Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res. Treat.* **146**, 287–297 (2014). doi: 10.1007/s10549-014-3019-2; pmid: 24929677
94. A. R. Grosso *et al*., Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. *eLife* **4**, e09214 (2015). doi: 10.7554/eLife.09214; pmid: 26575290
95. S. F. de Almeida, M. Carmo-Fonseca, Reciprocal regulatory links between cotranscriptional splicing and chromatin. *Semin. Cell Dev. Biol.* **32**, 2–10 (2014). doi: 10.1016/j.semcdb.2014.03.010; pmid: 24657193
96. M. E. Nemeroff, S. M. Barabino, Y. Li, W. Keller, R. M. Krug, Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Mol. Cell* **1**, 991–1000 (1998). doi: 10.1016/S1097-2765(00)80099-4; pmid: 9651582
97. A. J. Rutkowski *et al*., Widespread disruption of host transcription termination in HSV-1 infection. *Nat. Commun.* **6**, 7126 (2015). doi: 10.1038/ncomms8126; pmid: 25989971

ACKNOWLEDGMENTS

I thank my lab colleagues, especially M. Dye and H. Mischo, for advice on the text. I am indebted to H. Mischo for the illustrations. The expanding field of Pol II termination has generated a large literature. I apologize that space restricts its wider citation. My laboratory is supported by a Wellcome Trust Investigator Award (107928/Z/15/Z) and European Research Council Advanced Grant (339270).

10.1126/science.aad9926

RESEARCH ARTICLE SUMMARY

PROTEOMICS

Systems proteomics of liver mitochondria function

Evan G. Williams,* Yibo Wu,* Pooja Jha, Sébastien Dubuis, Peter Blattmann, Carmen A. Argmann, Sander M. Houten, Tiffany Amariuta, Witold Wolski, Nicola Zamboni, Ruedi Aebersold,† Johan Auwerx†

INTRODUCTION: Over the past two decades, continuous improvements in "omics" technologies have driven an ever-greater capacity to define the relationships between genetics, molecular pathways, and overall phenotypes. Despite this progress, the majority of genetic factors influencing complex traits remain unknown. This is exemplified by mitochondrial supercomplex assembly, a critical component of the electron transport chain, which remains poorly characterized. Recent advances in mass spectrometry have expanded the scope and reliability of proteomics and metabolomics measurements. These tools are now capable of identifying thousands of factors driving diverse molecular pathways, their mechanisms, and consequent phenotypes and thus substantially contribute toward the understanding of complex systems.

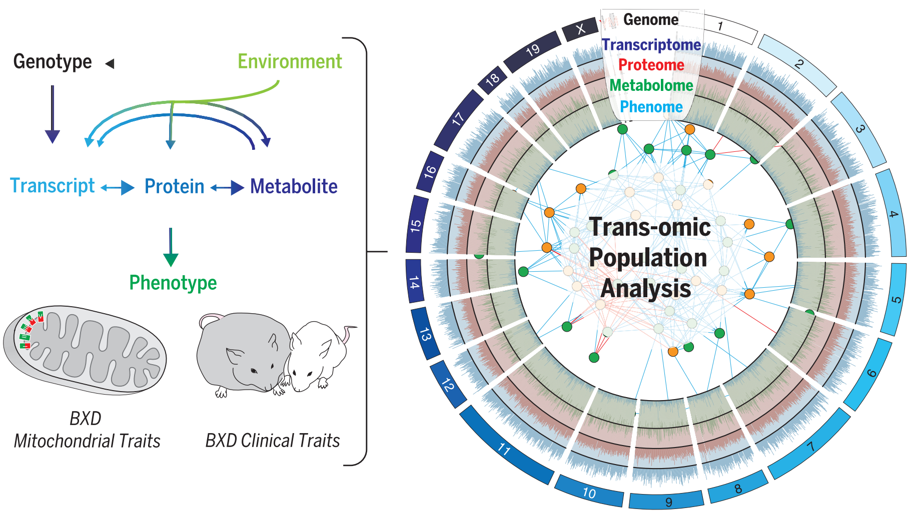
RATIONALE: Genome-wide association studies (GWAS) have revealed many causal loci associated with specific phenotypes, yet the identification of such genetic variants has been generally insufficient to elucidate the molecular mechanisms linking these genetic variants with specific phenotypes. A multitude of control mechanisms differentially affect the cellular concentrations of different classes of biomolecules. Therefore, the identification of the causal mechanisms underlying complex trait variation requires quantitative and comprehensive measurements of multiple layers of data—principally of transcripts, proteins, and metabolites and the integration of the resulting data. Recent technological developments now support such multiple layers of measurements with a high degree of reproducibility across diverse sample or patient cohorts. In this study, we applied a multilayered approach to analyze metabolic phenotypes associated with mitochondrial metabolism.

RESULTS: We profiled metabolic fitness in 386 individuals from 80 cohorts of the BXD mouse genetic reference population across two environmental states. Specifically, this extensive phenotyping program included the analysis of metabolism, mitochondrial function, and cardiovascular function. To understand the variation in these phenotypes, we quantified multiple, detailed layers of systems-scale measurements in the livers of the entire population: the transcriptome (25,136 transcripts), proteome (2622 proteins), and metabolome (981 metabolites). Together with full genomic coverage of the BXDs, these layers provide a comprehensive view on overall variances induced by genetics and environment regarding metabolic activity and mitochondrial function in the BXDs. Among the 2600 transcript-protein pairs identified, 85% of observed quantitative trait loci uniquely influenced either the transcript or protein level. The transomic integration of molecular data established multiple causal links between genotype and phenotype that could not be characterized by any individual data set. Examples include the link between D2HGDH protein and the metabolite D-2-hydroxyglutarate, the BCKDHA protein mapping to the gene *Bckdhb*, the identification of two isoforms of ECI2, and mapping mitochondrial supercomplex assembly to the protein COX7A2L. These respective measured variants in these mitochondrial proteins were in turn associated with varied complex metabolic phenotypes, such as heart rate, cholesterol synthesis, and branched-chain amino acid metabolism. Of note, our transomics approach clarified the contested role of COX7A2L in mitochondrial supercomplex formation and identified and validated *Echdc1* and *Mmab* as involved in the cholesterol pathway.

CONCLUSION: Overall, these findings indicate that data generated by next-generation proteomics and metabolomics techniques have reached a quality and scope to complement transcriptomics, genomics, and phenomics for transomic analyses of complex traits. Using mitochondria as a case in point, we show that the integrated analysis of these systems provides more insights into the emergence of the observed phenotypes than any layer can by itself, highlighting the complementarity of a multilayered approach. The increasing implementation of these omics technologies as complements, rather than as replacements, will together move us forward in the integrative analysis of complex traits. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aad0189



Model of the transomics analysis. A transomics approach was taken to analyze genetic and environmental variation in metabolic and mitochondrial phenotypes by measuring five distinct layers of biology in a diverse population of BXD mice. The combined analysis of all layers together provides additional information not yielded by any single omics approach.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: admin.auwerx@epfl.ch (J.A.); aebersold@imsb.biol.ethz.ch (R.A.)
Cite this article as E. G. Williams *et al.*, *Science* 352, aad0189 (2016). DOI: 10.1126/science.aad0189

RESEARCH ARTICLE

PROTEOMICS

Systems proteomics of liver mitochondria function

Evan G. Williams,[1*] Yibo Wu,[2*] Pooja Jha,[1] Sébastien Dubuis,[2] Peter Blattmann,[2] Carmen A. Argmann,[3] Sander M. Houten,[3] Tiffany Amariuta,[1] Witold Wolski,[2] Nicola Zamboni,[2] Ruedi Aebersold,[2,4]† Johan Auwerx[1]†

Recent improvements in quantitative proteomics approaches, including Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH-MS), permit reproducible large-scale protein measurements across diverse cohorts. Together with genomics, transcriptomics, and other technologies, transansomic data sets can be generated that permit detailed analyses across broad molecular interaction networks. Here, we examine mitochondrial links to liver metabolism through the genome, transcriptome, proteome, and metabolome of 386 individuals in the BXD mouse reference population. Several links were validated between genetic variants toward transcripts, proteins, metabolites, and phenotypes. Among these, sequence variants in *Cox7a2l* alter its protein's activity, which in turn leads to downstream differences in mitochondrial supercomplex formation. This data set demonstrates that the proteome can now be quantified comprehensively, serving as a key complement to transcriptomics, genomics, and metabolomics —a combination moving us forward in complex trait analysis.

Over the past two decades, continuous improvements in omics technologies have been driving an ever-greater capacity for quantifying relationships between genetics, the biochemical mechanisms of the cell, and overall phenotypes. Despite this progress, the majority of genetic factors determining complex trait heritability remain unknown (*1*). Recent advances in mass spectrometry (MS) (*2*–*4*) have expanded the scope and reliability of proteomic and metabolomic measurements. These developments in MS are permitting a leap forward in understanding complex biological systems by facilitating the accurate quantification of thousands of molecular factors involved in diverse cellular pathways—and, therefore, their mechanisms and consequent phenotypes (*5*). Thus far, the identification of causal genetic variants alone has been generally insufficient to characterize the underlying molecular mechanisms of action. Generating such models also requires quantitative measurements of additional layers of data, such as transcripts, proteins, and metabolites. As a multitude of control mechanisms differentially affect the cellular concentration of different classes of biomolecules, multilayered quantitative measurements on the same individuals can provide synergistic information about complex systems (*6*–*9*) [an approach also dubbed transomics or high-dimensional biology (*10*)].

In this study, we generated multilayered data sets to examine metabolism across 80 cohorts of the BXD genetic reference population (GRP). The BXDs are descended from C57BL/6J (B6) and DBA/2J (D2) and diverge for ~5 million sequence variants (*11*), similar to the number of common variants found within many human population groups (*12*). This population now consists of ~150 murine recombinant inbred strains with known variant susceptibility to major metabolic diseases such as diabetes (*13*, *14*). To date, detailed biochemical analyses have established validated links to phenotypes for a few dozen gene variants. These include links between sweet taste and *Tas1r3* (*15*), cadmium toxicity and *Slc39a8* (*16*), and hyperactivity and *Ahr* (*17*). We have previously reported that metabolic phenotypes in the BXDs are highly variable and that this variability is highly heritable (*14*). Here, we have analyzed 80 BXD cohorts (composed of 386 individuals) across a battery of metabolic tests such as ad libitum running-wheel access, maximal exercise capacity, and glucose tolerance. The mice were tested over a 29-week program where they were exposed to different environmental conditions of diet: chow diet (CD) (6% kcal of fat) or high-fat diet (HFD) (60% kcal of fat). To understand the molecular basis behind the observed phenotypic variance, we quantified detailed layers of systems-scale molecular measurements in the livers of the entire population: the transcriptome (25,136 transcripts), the proteome (2622 proteins), and the metabolome (981 metabolites). Together with full coverage of genetic variants in the BXDs (*18*), these omics data sets provide a comprehensive platform for deconstructing the factors behind variation in clinical metabolic phenotypes. In all layers of data, trait variation could be attributed to the causal genetic loci through quantitative trait locus (QTL) analysis. These layers build on our previous research in this population (*19*), in which selected reaction monitoring (SRM) was used to quantify 192 proteins and targeted metabolomics approaches were used to quantify 39 metabolites in the serum and 2 metabolites in the liver. This previous study both shaped our bioinformatics procedures for transomic data sets and provided positive controls for the experimental Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH-MS) proteomics and for multilayered pathway analysis. For example, of the 13 genes with cis-pQTLs (protein) of the 192 proteins measured by SRM, 11 were identified in SWATH (all except AKR7A2 and ABCB8), and 10 of these 11 also mapped to cis-pQTLs in the independent SWATH data set.

By applying transomic analyses in these data sets, we observed that the levels of all four proteins composing the branched-chain ketoacid dehydrogenase (BCKD) complex in the mitochondria are, in the BXDs, tied to genetic variants in a single gene, *Bckdhb*. Similarly, a causal mechanistic link was observed between the D-2-hydroxyglutarate dehydrogenase (D2HGDH) protein and the metabolite D-2-hydroxyglutarate (D2HG), which in turn is linked with similar phenotypes as for humans with deficiencies in this protein, including cardiomyopathy and problems with motor control (*20*). Furthermore, the broad proteomics data set allowed us to identify two isoforms of the protein ECI2 that were not predicted by eQTL (expressed transcript) analysis. We examined several broad pathways in energy metabolism using a transomics approach, including lipid storage/transport, cholesterol synthesis, and the electron transport chain (ETC), all of which exhibited high levels of genetic variation at the transcript, protein, and metabolite levels. This analysis highlighted COX7A2L from the ETC—the only one of 67 quantified proteins in the ETC with consistent cis-pQTLs. Further experiments showed that this variation in protein leads to strikingly differential formation of ETC supercomplexes (SCs). In all cases, the integrated analysis of multiple omics layers provided more insight into mechanistic networks than could be gleaned from any layer by itself, highlighting the complementarity of a multilayered approach.

High-dimensional reconstruction of complex metabolic traits

To identify new genetic relationships and molecular mechanisms influencing metabolism in the BXDs, we designed an analytical pipeline to measure and combine quantitative data from five omics layers across variable environmental states: genomics, transcriptomics, proteomics, metabolomics, and phenomics (Fig. 1A). The 29-week phenotyping program includes body weight, indirect calorimetry, voluntary exercise, maximal oxygen consumption (VO_{2max}), an oral glucose tolerance test, and spontaneous activity (Fig. 1B). All traits vary significantly due to genetic, environmental,

[1]Laboratory of Integrative and Systems Physiology, Interfaculty Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, CH-1015, Switzerland. [2]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, CH-8093, Switzerland. [3]Department of Genetics and Genomic Sciences and Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, Box 1498, New York, NY 10029, USA. [4]Faculty of Science, University of Zurich, CH-8057, Switzerland.
*These authors contributed equally to this work.
†Corresponding author. Email: admin.auwerx@epfl.ch (J.A.); aebersold@imsb.biol.ethz.ch (R.A.)

and/or gene-by-environment (G×E) factors, including key traits such as body weight (Fig. 1C) and glucose response during an oral glucose tolerance test (Fig. 1D). At the end of the program, liver samples were used for multilayered omics analyses to serve as the platform for determining the providence and mechanism of metabolic variants across the population. Together, these data support approaches driven by prior knowledge—e.g., examining the relationships between known transcriptional and proteomic gene networks with related phenotypes—as well as data-driven approaches—e.g., the de novo identification of genes that are involved in regulation of metabolic phenotypes.

Before delving into multilayered data sets for the analysis of complex molecular networks, we assessed the quality and significance of each omics layer individually. For phenotypes, heritability was calculated for all traits within dietary groups (green and black bars, Fig. 1E), then across all cohorts combined (red bars). The known environmental factor, diet, furthermore allowed us to calculate the independent environmental effect (blue) and the strain-dependent G×E influence (yellow). As expected, HFD feeding strikingly worsens parameters of metabolic health in most strains, particularly for traits such as body weight, glucose response, and running capacity (Fig. 1F). However, we observed a tremendous range in HFD response among BXD strains: The body weight in some strains is unchanged (e.g., BXD68), whereas others nearly double in size (BXD44) (Fig. 1C). Next, we examined the transcriptome data (Affymetrix Mouse Gene 1.0 ST microarrays), in which 25,099 annotated probe sets were quantified. Of these, 21,970 were designated as protein coding, with the unaligned transcripts corresponding to noncoding genes such as open reading frames and putative Riken cDNAs or genes that only have unreviewed UniProt identifiers. When transcript levels were compared across diets, many of those most strongly modulated conform to expectations from the literature—e.g., *Fosl1* is down-regulated in HFD (*21*) and *Pparg* is up-regulated in HFD (*22*) (Fig. 1G)—although the cause of variation in other top genes is less clear (e.g., *Thnsl1*). Likewise for metabolites; some were observed to be strongly affected by dietary and genetic factors in both diets [e.g., farnesyl pyrophosphate (FPP)], and others fluctuated in only one diet (e.g., allyl isothiocyanate)

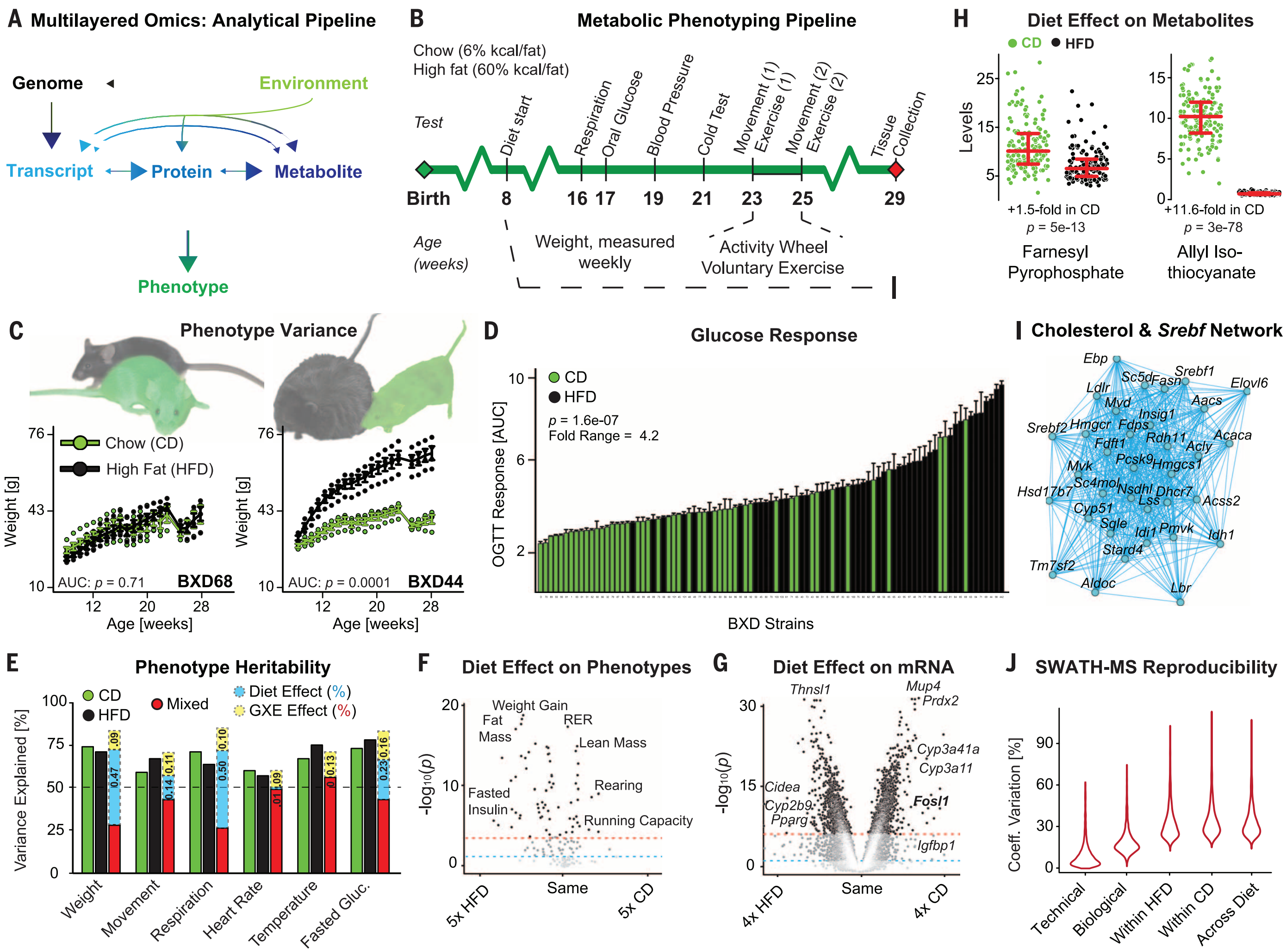


Fig. 1. Overview and validation of omics layers. (**A**) General model of the multilayered approach. Arrows indicate causality between metabolic layers. HFD should not modify DNA, although other environmental factors can (i.e., mutagens). (**B**) Phenotyping pipeline for all individuals. (See the methods section for details on each experiment.) (**C**) Body weight in two strains of BXD for both diets over the full phenotyping experiment. (**D**) Area under the curve (AUC) of glucose excursion during a 3-hour oral glucose tolerance test for all cohorts. Bars represent mean ±SEM. (**E**) Heritability for several phenotypes, calculated by one-way (CD/HFD) or two-way (Mixed) analysis of variance. Some traits are affected by diet (weight and fasted glucose), others are not (heart rate and body temperature), and G×E contributions vary. (**F**) Volcano plot of diet effect on clinical phenotypes. (**G**) Volcano plot of diet effect on all transcripts. (**H**) Dot plot showing two example hepatic metabolites affected by diet. (**I**) An enriched Spearman correlation transcript network using the cholesterol biosynthesis and SREBF targets gene set. Edges indicate $P \leq 0.001$. All correlations are positive. (**J**) Error in SWATH measurements due to different factors: technical (median CV = 6.5%), biological (CV = 17.0%), across strain (within diet) (CV = 29.6% HFD, 31.4% CD), and across all measurements (CV = 30.8%). Reported P values between diets (panels C–D, F–H) are all for Welch's t-test.

(Fig. 1H). Network analysis of large and metabolically relevant gene sets in these omics layers, such as cholesterol biosynthesis (Fig. 1I), showed high levels of enrichment in transcript covariation compared with noise, indicating that the variant transcript levels are functionally linked and physiologically relevant (described in more detail later). We then assessed the data generated by SWATH-MS proteomics (3). As SWATH-MS is an emerging technology, we performed several additional checks, including the technical error, the biological error within cohort, and the errors within diet and across all samples (Fig. 1J). Reproducibility was excellent, with median technical error being ~8% of overall variation in protein levels. Similarly, variation within biological replicates was much lower than variation across the genotypes or dietary conditions (Fig. 1J).

Next, we examined transcript-protein relationships. Of the 2622 unique proteins quantified in all cohorts, 2600 aligned to measured transcripts. Spearman correlation analysis was performed for all pairs in both diets independently. The data indicated that 1004 transcript-protein pairs correlate at nominal significance in CD (raw $P < 0.05$) (Fig. 2A) and 938 in HFD cohorts. Of these, 637 pairs (~25%) correlated at least nominally significantly in both diets (Fig. 2B, green, purple, and red points). This moderate—although still highly significant—correlation between genes'

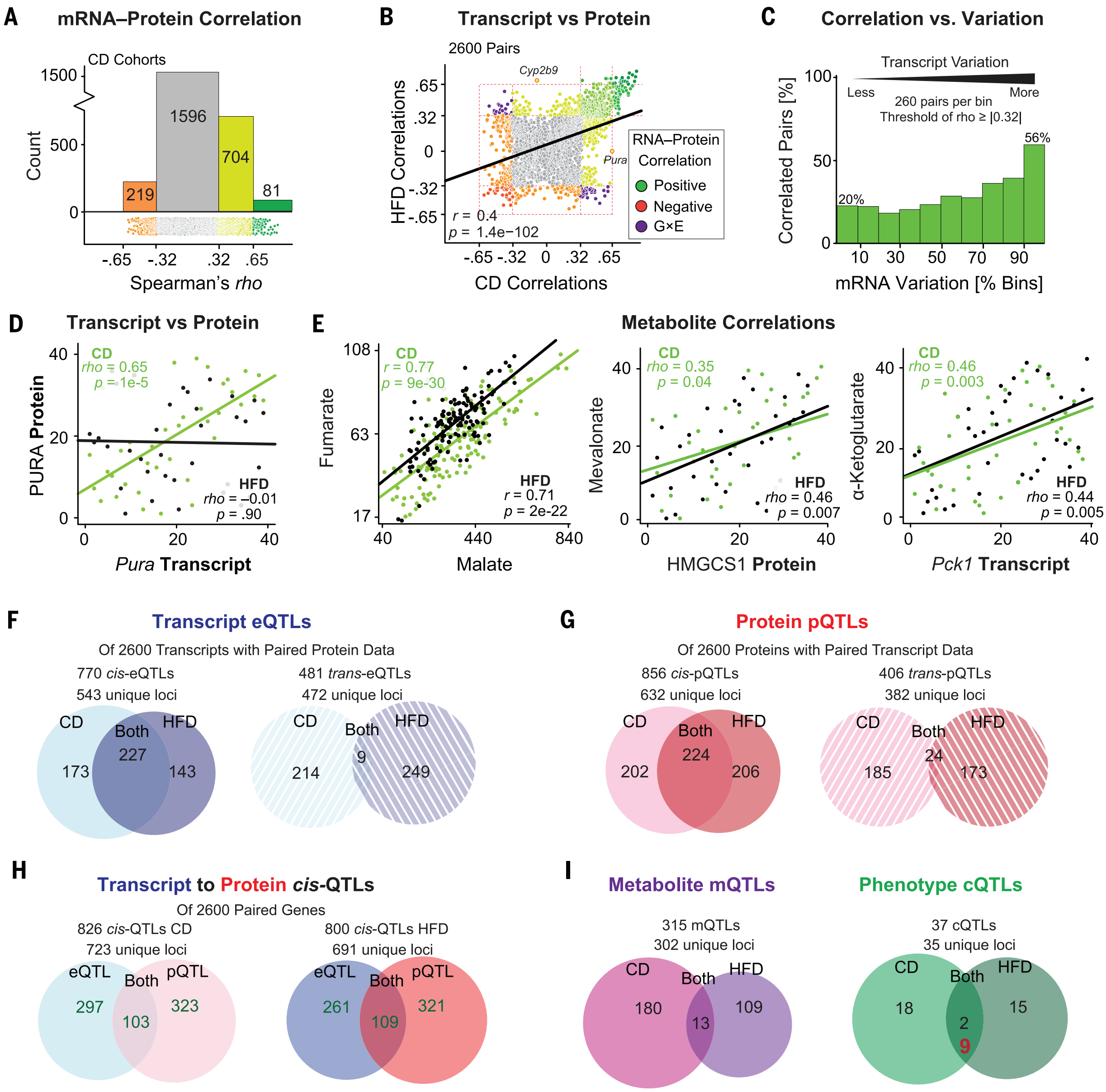


Fig. 2. Multilayer analysis of associations and causality. (**A**) Histogram of 2600 transcript-protein pair Spearman correlations in CD. $\rho = 0.32$ corresponds to a nominal $P < 0.05$. $\rho = 0.65$ corresponds to Bonferroni-corrected significance. (**B**) Correlation plot of transcript-peptide Spearman correlation coefficients in CD against HFD. (**C**) Transcript-protein correlation prevalence in CD cohorts, binned by transcript variation. Among the top 10% most variable transcripts (260 pairs), 56% of pairs correlate, in contrast to only 20% of pairs in the lowest bin. Nominal significance cutoffs are used, so ~5% of matches in each bin are false positives. (**D**) The transcript *Pura* correlates significantly with its protein in CD but not in HFD. (**E**) Malate and fumarate, two adjacent metabolites in the TCA cycle, correlate strongly. Several other cross-layer correlations are observed between metabolites and their adjacent enzymes in major metabolic pathways. (**F**) Venn diagram and count of all cis- and trans-eQTLs across diets for the 2600 transcripts with matching protein measurements. (**G**) Venn diagram and count of all cis- and trans-pQTLs for the same 2600 proteins. (**H**) Overlap between cis-eQTLs and cis-pQTLs in both diets. Fifty-nine genes map to cis-QTLs in all four data sets (intersection not shown). (**I**) Venn diagram of all mQTLs and cQTLs in both diets. In red for cQTLs: overlapping cQTLs that are genome-wide significant in one diet (LRS ≥ 18) and locally significant in the other (LRS ≥ 12).

transcripts and proteins is in line with previous population studies, which examined smaller numbers (hundreds) of such pairs (19, 23). Variation in any given transcript or protein's expression within a single tissue in normal population data sets can be highly variable. Among the 2600 transcript-protein pairs, 90% of transcripts vary between 1.4- and 4.0-fold across all samples. The magnitude of this variance strongly influences cross-layer correlation. When the 2600 paired transcripts are binned into 10 groups based on expression range across CD cohorts, the highest bin has 260 transcripts with ≥2.8-fold range, whereas the lowest bin has 260 transcripts with ≤1.4-fold range. Correspondingly, 56% of transcript-protein pairs correlate at $\rho \geq |0.32|$ (raw $P \leq 0.05$) in the top bin, versus only 20% in the lowest bin (Fig. 2C). Fundamentally, a protein cannot exist if there is no corresponding transcript; thus in some sense, 100% of transcript-protein pairs could be considered correlated. Conversely, we observe that only ~15 to 20% of transcript-protein pairs are reactive to small changes in the other's expression (i.e., the lower bins) (Fig. 2C).

As with phenotypes and transcripts, G×E effects were observed in transcript-protein correlations, such as for the several dozen transcript-protein pairs whose correlation segregated depending on diet (Fig. 2B, purple dots). For the 137 most significant correlations—those that met the Bonferroni-corrected significance threshold (corresponding to $\rho = \pm 0.65$) in at least one diet—135 correlated at least nominally significantly ($P < 0.05$) in the other diet with the same directionality. For the other two, *Cyp2b9* and *Pura*, strong correlation was observed in one diet and no correlation in the other (e.g., Fig. 2D). This frequent discrepancy between variation in transcript and protein levels indicates that reanalyzing metabolic pathways using more comprehensive proteomic coverage can identify unknown biological mechanisms. Last, we examined the metabolomics layer. Here, metabolite signatures of 979 unique mass-to-charge ratios (m/z) were measured in 357 liver samples using a time-of-flight MS approach (24). These 979 features were then aligned to specific chemical signatures using previously assembled reference libraries (24). Initial data quality checks were performed by analyzing successive metabolites within pathways, such as the tricarboxylic acid (TCA) cycle and glycolysis. Clear connections were frequently observed between consecutive metabolites of a given pathway, such as between malate and fumarate (Fig. 2E). Although metabolites do not fit as neatly into the direct relationships of gene-transcript-protein, the measurement quality and the physiological relevance of metabolite variation may be examined through relationships between the transcript and protein levels of different enzymes with their up- or downstream metabolites. In this analysis, cross-dimensional correlations between known factors were frequently observed, including for cholesterol biosynthesis, glycolysis, and the TCA cycle (25) (Fig. 2E). Together, these validation steps confirm the general data quality and reliability and the potential of a multilayered analytical approach.

Metabolic relationships to multilayered data

We next sought to identify causal genetic mechanisms that determine molecular expression levels

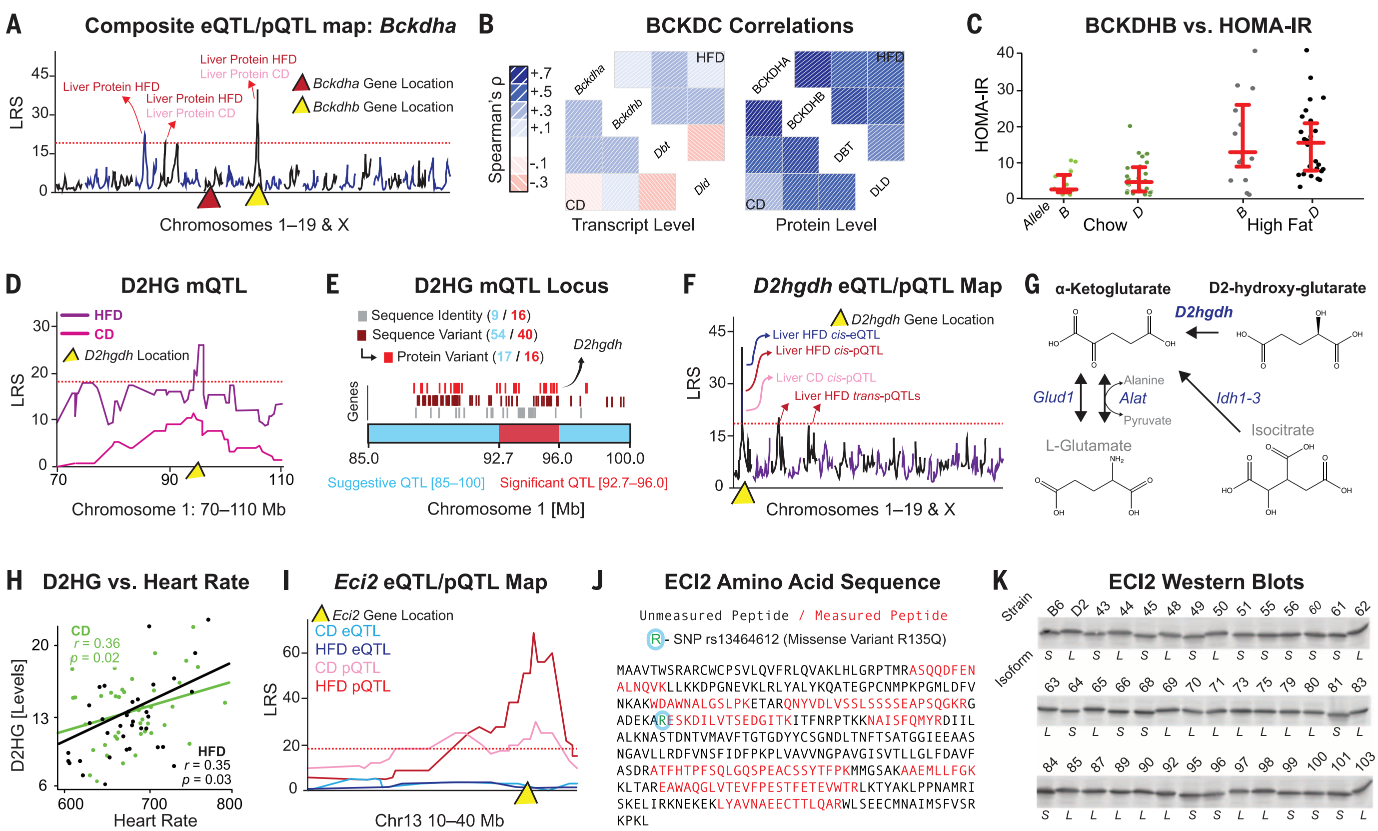


Fig. 3. Identifying the QTGs and causal mechanisms driving QTLs. (**A**) Combined QTL map of *Bckdha* transcript and protein in both diets. Significant trans-pQTLs map to *Bckdhb* (yellow triangle) on chromosome 9, whereas no cis-QTLs map to *Bckdha* on chromosome 7 (red triangle). (**B**) Spearman correlation matrices of the four subunits of the BCKDC at the transcript or protein level in both diets. (**C**) Homeostatic model assessment for insulin resistance (HOMA-IR) is significantly increased in HFD cohorts compared with CD ($P = 2 \times 10^{-6}$, Welch's t test), but no association is seen between *Bckdhb* allele and HOMA-IR in either dietary cohort. (**D**) D2HG maps significantly to chromosome 1 in the HFD cohort. (**E**) This locus contains 56 genes, of which 16 have a major genetic variants variable, including *D2hgdh*. (**F**) Composite eQTL and pQTL map for *D2hgdh*. The protein maps as a cis-pQTL in both diets to the same chromosome 1 locus, whereas only the HFD transcript levels map to a cis-eQTL. (**G**) *D2hgdh* drives one of several pathways generating α-ketoglutarate. (**H**) D2HG is positively associated with heart rate in both diets in a Pearson correlation. (**I**) *Eci2* exhibits no cis-eQTLs, but yields significant cis-pQTLs in both diets. (**J**) Peptide sequence analysis of ECI2, with the nine measured peptides and the single missense mutation highlighted. (**K**) ECI2 Western blots show two distinct molecular weight bands depending on the BXD genotype.

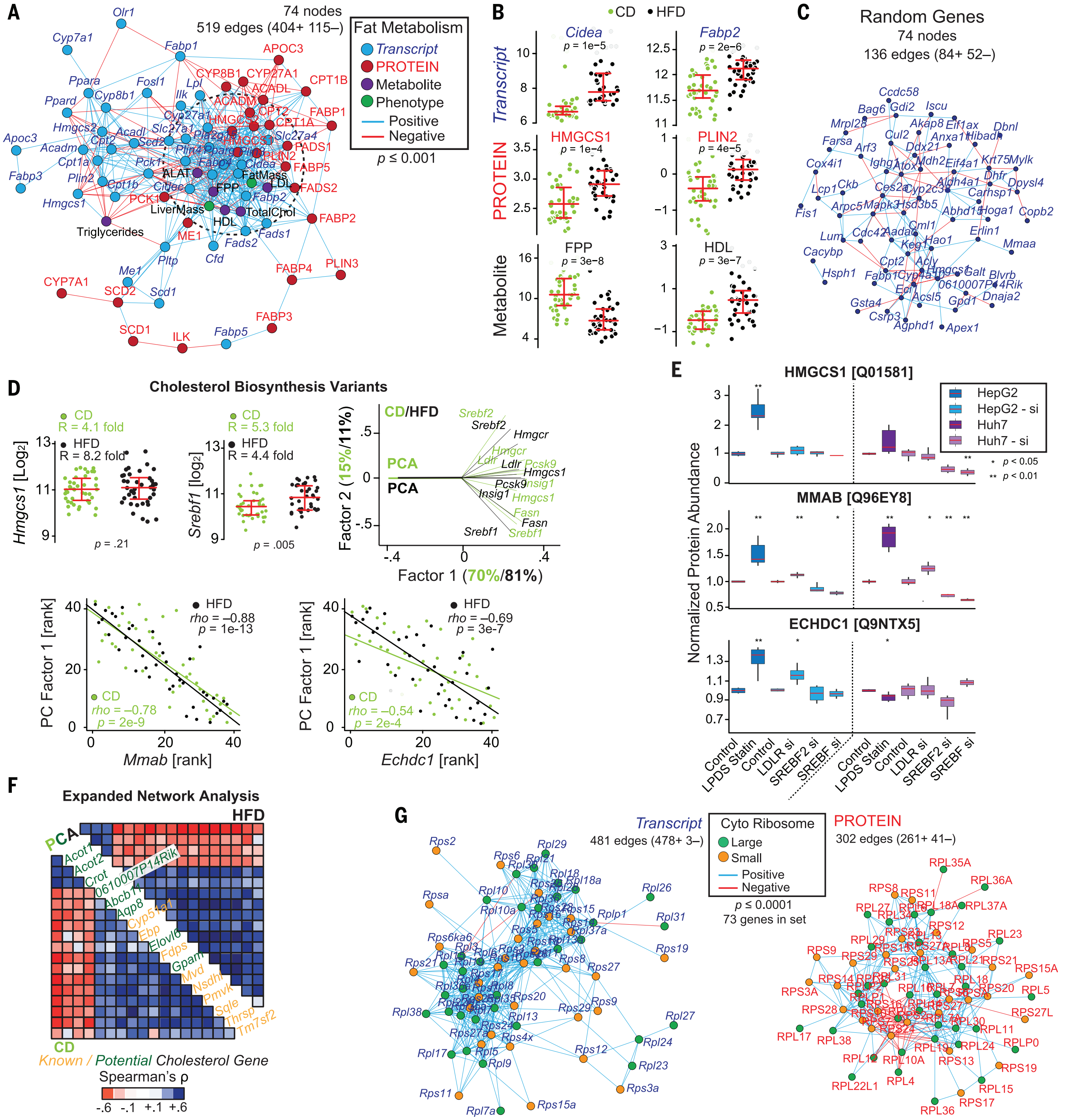


Fig. 4. Network analysis. (**A**) Spearman correlation network showing mixed genes involved in fat metabolism at the transcript and protein level, along with key metabolites and phenotypes. The dashed circle represents the core enriched gene set. Edges are significant at $P < 0.001$ for a positive (blue) or negative (red) correlation. (**B**) Diet-dependent expression of key genes and metabolites involved in fat metabolism; P values are for Welch's t test. (**C**) A Spearman correlation network of 74 transcripts taken at random from the list of 2600 genes measured at the transcript and protein level, using the same network analysis. Edge counts correspond to the level expected from noise. (**D**) (Left) *Hmgcs1* and *Srebf1*, as well as other transcripts and proteins in the cholesterol biosynthesis pathway, are highly variable in the BXDs. (Right) PCA of a set of eight cholesterol biosynthesis genes shows that their variances are highly explained by a single factor. (Bottom) Two candidate cholesterol genes, *Mmab* and *Echdc1*, which correlate with PC1 in both diets. (**E**) In vitro validation of HMGCS1, along with two proteins not known to be involved in cholesterol metabolism—MMAB and ECHDC1—which clustered with known cholesterol genes. MMAB and ECHDC1 both respond like HMGCS1 to lipid-deficient serum and statin treatment or to knockdown of LDLR, SREBF2, or SREBF1/2, suggesting that they are indeed involved in cholesterol metabolism. (**F**) Unbiased Spearman correlation matrices of the first PC in CD (bottom left) and HFD (top right) conditions with other genes turned up many known cholesterol-regulatory genes (orange) as well as new candidates (green). (**G**) Transcript and protein networks for the 73 genes with paired transcript-protein data in the cytosolic ribosome complex. Both were highly enriched, although with tighter coregulation at the transcriptional level. Edges represent Spearman correlations with $P < 0.0001$.

of any of the omics layers through QTL analysis. The cohorts in both dietary states were analyzed for eQTLs, pQTLs, mQTLs (metabolite), and cQTLs (clinical phenotype). Across all 2600 transcripts for which we generated associated protein data, 543 genes mapped to 770 significant cis-eQTLs and 472 genes mapped to 481 distinct trans-eQTLs (Fig. 2F) (QTLs detected in both diets are considered twice). Of the 543 genes with cis-eQTLs, 227 mapped consistently across both diets (41%), whereas trans-eQTLs rarely overlapped (2% overlap detected, with ~0.2% overlap expected by chance) (Fig. 2F). At the protein level, 632 distinct genes mapped to 856 cis-pQTLs, and 382 genes mapped to 406 distinct trans-pQTLs (Fig. 2G). Across diets, we observed that 35% of cis-pQTLs mapped in both diets, similar to the ratio for transcripts. Consistent trans-pQTLs were still quite rare, albeit more common than for transcripts (~6%). Roughly 4% of examined genes mapped as cis-eQTLs and cis-pQTLs in at least one diet (103 or 109 versus 2600), whereas of genes with a cis-QTL of any sort, roughly 85% were unique to the transcript or protein level (103 of 826 for CD and 109 of 800 for HFD were shared cis-QTLs) (Fig. 2H). Fifty-nine genes (2.2%) mapped to cis-QTLs in both layers and both diets. For the metabolite layer, 315 significant mQTLs (LRS ≥ 18) were detected, of which 13 mapped consistently in both dietary states (~4%). For phenotypes, we calculated 37 significant cQTLs, of which 2 mapped significantly to the same locus in both diets: movement activity [caused by *Ahr* (*17*)] and heart rate (causal gene unknown). To increase the scope of QTLs consistent across diets, one dietary data set may also be used to identify a hypothesis at system-wide significance (LRS ≥ 18), and the other diet may be used to test the hypothesis at locus-specific significance (LRS ≥ 12). In doing so, an additional 7 cQTLs are observed as consistent in both diets (Fig. 2I, red number).

Solving QTLs: Finding the quantitative trait gene

For cis-QTLs, the causal factors can be quickly identified: With few exceptions, they will be driven by variants within the gene itself or immediately adjacent. For trans-QTLs, mQTLs, and cQTLs, the identification of the causal quantitative trait gene (QTG) is challenging due to the width of the QTLs. In the BXDs, QTLs calculated using 40 strains are typically 2 to 8 Mb wide, with an average of 10 genes per Mb (*14*). We first examined the 24 genes with trans-pQTLs that were observed in both dietary cohorts to search for QTGs and downstream effects. One of these 24 genes is *Bckdha*, which encodes the E1 alpha polypeptide of the BCKD complex. The BCKDHA protein levels map to a trans-pQTL on chromosome 9 in both diets, whereas no such trans-eQTL is observed for the *Bckdha* transcript (Fig. 3A). Strikingly, this locus contains the E1 beta polypeptide, *Bckdhb*, which itself has cis-pQTLs and cis-eQTLs in the BXDs (*19*). *Bckdha* and *Bckdhb* encode the E1 subunit of the BCKD complex, which, together with the E2 (*Dbt*) and E3 (*Dld*) subunits, regulates the breakdown of branched-chain amino acids (BCAAs). Variants in either E1 subunit can lead to an inborn error of metabolism called type 1 maple syrup urine disease (MSUD) (*26*), the biochemical features of which some strains of the BXDs are known to exhibit (*19*). Notably, the transcript levels of these four genes coding for the complex components have no strong association, whereas the expression levels of the proteins are strongly coupled (Fig. 3B). That is, genetic variants in *Bckdhb* are causal for protein level variation in BCKDHA and BCKDHB, which in turn correlate strongly positively with dihydrolipoamide branched chain transacylase E2 (DBT) and dihydrolipoamide dehydrogenase (DLD) levels in both diets: the levels of all four proteins are linked. Conversely, although *Bckdhb* gene variants cause transcriptional changes in *Bckdhb* mRNA (a cis-eQTL), there is no effect on *Bckdha* mRNA, nor any correlation with *Dbt* or *Dld*. Although a significant difference in the BCAA/alanine ratio across the BXDs has been observed and linked to the *Bckdhb* allele (*19*), we observe no association with other metabolic hallmarks of MSUD, such as insulin or glucose levels (Fig. 3C). However, such a link may only be observed in more exacerbated states, such as in cohorts fed diets high in BCAAs, as suggested by prior literature (*27*). This example highlights the importance of examining protein levels in the diagnosis and elucidation of other metabolic diseases and underlines the possibility of using the BXDs as a MSUD model. Additionally, this finding highlights the possibility that the expression levels of proteins within multigene complexes may be more tightly coregulated than are their transcripts (e.g., Fig. 3B).

We then aimed to identify candidate QTGs behind any of the 302 distinct mQTLs using orthogonal pQTL data. All 856 genes with cis-pQTLs were compared against these mQTLs, with the hypothesis that genes with cis-pQTLs are more likely to be causal for other QTLs mapping to their locus (*28*). This process highlighted several potential pQTL/mQTL links. In particular, D2HG maps to a locus on chromosome 1 containing 56 genes (Fig. 3, D and E). Among these 56 genes is D-2-hydroxyglutarate dehydrogenase (*D2hgdh*), which maps consistently to cis-pQTLs (Fig. 3F) and correlates negatively with the upstream metabolite D2HG ($\rho = -0.37$ and $\rho = -0.48$ in CD and HFD, respectively). Although the other 55 genes in this locus could contribute to this mQTL, *D2hgdh* is known to convert D2HG to α-ketoglutarate in the mitochondria (*29*) (Fig. 3G). In humans, variants in *D2HGDH* have been linked to severe disease traits such as cardiomyopathy and motor difficulties (*20*). In the BXDs, we observe moderate but consistent connections between D2HG and cardiovascular phenotypes such as heart rate (Fig. 3H) and exercise capacity (moderate negative correlation), indicating that some mild aspects of the disease phenotype may manifest in this population, even under nonstressed conditions.

The SWATH proteomics analysis is also able to identify nonsynonymous sequence variants across the BXDs, which are detected as peptide-specific cis-pQTLs [cis-peptide(pep)QTLs]. To demonstrate this, we highlight enoyl-CoA delta isomerase 2 (*Eci2*), a mitochondrial enzyme involved in fatty acid oxidation. Nine distinct peptides were quantified for ECI2, of which one displayed a striking cis-pepQTL in both diets. Interestingly, there are no cis-eQTLs for ECI2 at the gene or exon level (Fig. 3I). Variant analysis revealed a nonsynonymous change (R135Q, rs13464612) adjacent to this peptide that is predicted to abolish the trypsin cleavage site (Fig. 3J). Furthermore, this cis-mapping variant tracks with a small change ECI2 migration in SDS–polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 3K), indicating a change in the protein that is not observed at the transcript level. Notably, these analyses highlight the ability to detect putative protein isoforms—an additional source of molecular variation underlying complex phenotypes.

High-dimensional metabolic networks

As shown above, large, multilayered quantitative omics data sets can be used to map and solve QTLs at high throughput. Perhaps the more unique characteristic of more comprehensive measurement techniques, however, is that the resulting data may be used to model extended pathways or functional networks with dozens of proteins and metabolites acting in tandem. To examine this possibility in the BXDs, we performed ontology enrichment analysis on 226 KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways (*30*) to identify which other metabolic pathways are best covered by the proteomic and metabolomics data and may benefit from a transomic approach. Among the most enriched pathways are those involved in fatty acid metabolism and storage, such as the peroxisome proliferator-activated receptor (PPAR) pathway. These sets of genes and pathways are furthermore known to be variable in the BXDs and lead to overt differences in fatty liver and metabolic changes in the liver, including metabolic signatures of liver stress such as increased alanine transaminase (ALAT) and disease phenotypes such as fatty liver disease (*21*). Of the 41 genes in this pathway, 25 were measured at the protein level, along with a handful of relevant metabolites [e.g., FPP, ALAT, and high-density lipoprotein (HDL)]. Interestingly, although the transcripts and proteins for any single gene did not strikingly correlate (Fig. 4A), significant enrichment of correlations between transcripts and proteins was observed in other parts of the pathway (Fig. 4A, dashed circle), indicating interactions between these two layers and between related metabolic pathways. In turn, the variations in these transcripts and proteins contributed to proximal metabolic changes [e.g., low-density lipoprotein (LDL), HDL, and FPP levels] and to related global phenotypes such as total fat and liver mass. This network is also highly responsive to diet, with clear differences between CD and HFD cohorts in key genes and metabolites (Fig. 4B). While these findings are expected (*21*), they further emphasize the reliability of cross-layer omics analysis and furthermore provide better networks than those expected using random gene sets (Fig. 4C).

Within the fat metabolism gene set, we observed particularly high levels of variability in key genes in cholesterol biosynthesis such as *Srebf1*

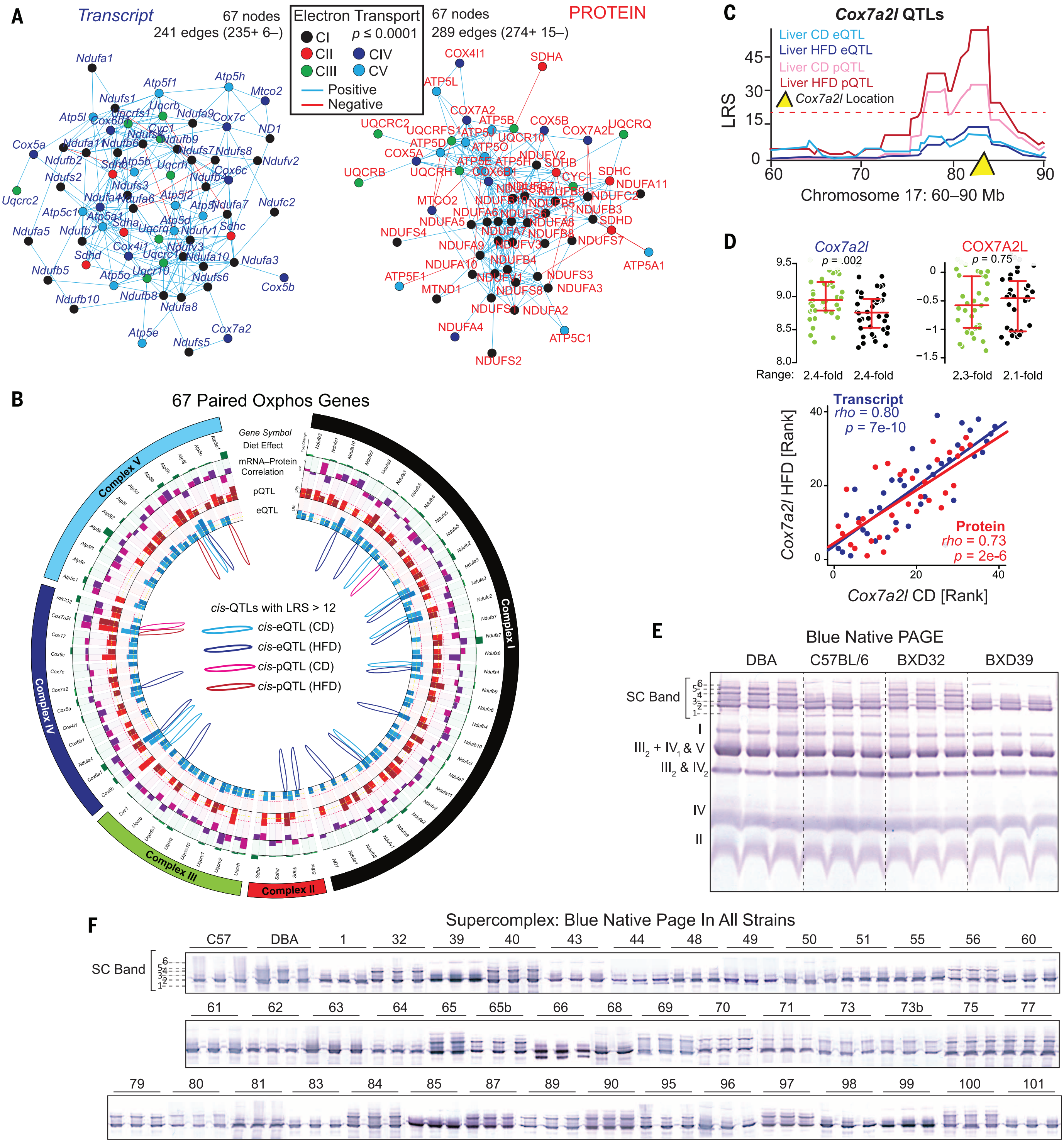


Fig. 5. Variable mitochondrial phenotypes in the BXDs. (**A**) The oxphos protein Spearman correlation network is somewhat more tightly coregulated than the transcript network. In particular, CI proteins cluster more tightly than CI transcripts (black nodes). (**B**) Circos plot of 67 ETC genes. Green bar ring: effect of diet, relative change between medians. Light green: transcript; dark green: protein. Purple bar ring: correlation between transcript and protein in CD (light purple) or HFD (dark purple). Red bar ring: LRS of peak pQTLs in CD (light red) and HFD (dark red). Blue bar ring: LRS of peak eQTLs in CD (light blue) and HFD (dark blue). Inside: drawing of significant cis-QTLs (LRS ≥ 12). Significant trans-QTLs (LRS ≥ 18) are not drawn. (**C**) Diet-consistent cis-pQTLs were observed only for COX7A2L, which does not map to significant cis-eQTLs. (**D**) (Top) The *Cox7a2l* transcript is affected by diet, whereas both transcript and protein are highly variable across genotype. (Bottom) Expression is consistent across diets within the transcript and protein level, despite the presence of dietary effect in mRNA and its absence in protein. (**E**) BN-PAGE for four strains with three biological replicates. Individual complexes are labeled. Several distinct upper SC bands are observed, labeled initially as 1 through 6. (**F**) Upper SCs for all CD cohorts (several independent gels are aligned and spliced together). SCs were quantified in binary fashion by presence (+1) or absence (0) of a particular band.

and *Hmgcs1* (Fig. 4D, left). Furthermore, key *Srebf* target genes, including *Hmgcr, Pcsk9, Insig1,* and *Fasn*, all clustered tightly in principal components analysis (PCA), with the first principal component of eight core genes explaining 70 and 81% of the total variation in CD and HFD, respectively (Fig. 4D, right). Correlation analyses of the first principal components across the transcriptome and proteome data sets yielded several hits that correlate strongly and independently in both CD and HFD cohorts. These genes included several not known to be involved in cholesterol biosynthesis, such as *Mmab* and *Echdc1* (Fig. 4D, bottom), as well as a number of genes involved in cholesterol that were not included in input, such as *Fdps*, *Mvd*, and *Dhcr7*. We selected two hits, *Mmab* and *Echdc1*, which are not well reported in cholesterol literature, for in vitro validation. In HepG2 and Huh7 cell lines, we examined the effect of cells in lipoprotein-deficient serum (LPDS) treated with statins and the effect of *Ldlr*, *Srebf2*, or *Srebf1/2* small interfering RNA (siRNA) knockdown—conditions that modulate different aspects of the cholesterol biosynthesis pathway (*31*–*33*). We observed that MMAB and ECHDC1 proteins are modulated strongly by LPDS + statin treatment and behave similarly (although not identically) to HMGCS1, an established *Srebp2* target protein and a key regulatory gene in cholesterol synthesis. Similar genome-wide analyses indicated additional candidate cholesterol-related genes, such as *Aqp8*, *0610007P14Rik*, and *Gpam* (Fig. 4F). Some of these genes have been identified in previous cholesterol genome-wide association studies, such as for *Mmab* with blood HDL levels (*34*, *35*), whereas other candidates are likely involved in tangential pathways (e.g., *Acot1* and *Acot2* are involved in lipid metabolism).

We next examined gene sets that form large and cohesive protein structures, such as the ribosome. Pairwise covariance analysis of transcripts coding for ribosomal proteins have previously been shown to form a tightly connected network (*36*). To extend this analysis to the protein level, we systematically generated Pearson correlation networks for the 73 genes in the ribosome family that were measured at both the transcript and protein level in all samples. As for the transcripts, the proteins clustered into enriched correlation networks (Fig. 4G). That the ribosomal genes are coregulated is expected, but it illustrates that the data are reliable enough to reveal components of functional modules and thus to support systems analyses.

Following these proofs of concept, we determined which metabolic pathways were most comprehensively covered in the multilayered data sets and triaged them for further analysis, with particular focus on mitochondria-related sets. Among the groups with the most complete protein coverage was the oxidative phosphorylation gene set (oxphos, hsa00190). This gene set is defined by 133 genes, of which 70 were quantified at the protein level. Of these, 67 were also quantified at the transcript level—all except the mitochondrial-encoded *ND4*, *ND5*, and *ATP8*. We performed network analyses on these data and observed highly significant positive correlation networks (Fig. 5A), with the protein network somewhat more enriched than the transcript network. This layer-specific difference in network structure is perhaps logical, given that the proteins are bound together in stoichiometry in their functional complexes, whereas the equivalent transcripts are not. The oxphos network was not strongly affected by diet, with only four proteins and 16 transcripts being variable at the permissive cutoff of $P < 0.01$. (Zero proteins and four transcripts—*Ndufb5*, *Cox7a2*, *Atp5b*, and *Ndufa7*—are significantly influenced at adjusted $P < 0.05$.) Furthermore, observed dietary differences at the transcript level did not reliably indicate any similar change in protein levels (Fig. 5B, outermost green bands.)

Consequences of oxphos variation

Given the relatively small effect of diet on transcript and protein levels, surprisingly few QTLs were observed consistently across diet: only six transcripts and one protein (Fig. 5B). For the only such protein, COX7A2L, no corresponding cis-eQTLs were observed (Fig. 5C). Interestingly, the *Cox7a2l* transcript exhibited significantly different expression in response to diet, whereas the protein levels were unaffected (Fig. 5D, top). The transcript and protein levels were highly correlated across dietary cohorts, suggesting a strong genetic influence on both *Cox7a2l* and COX7A2L levels, independent of the observed dietary influence on the transcript (Fig. 5D, bottom). Because mitochondrial transcript and protein networks are highly variable and coregulated in the BXDs (Fig. 5A), we hypothesized that these may be associated with clear differences in mitochondrial structures and phenotypes. To broadly test this idea, we performed blue native (BN)–PAGE analysis on isolated mitochondria from three biological replicates across all strains, using the same liver samples as before. Mitochondrial complex levels and formations varied across the BXDs (Fig. 5E), with particularly striking differences in SC formations (Fig. 5F, all strains). The differences in SC patterns across strains, and the consistency within strains, indicate that complex and multifactorial genetic interactions are driving the mitochondrial effects, at least in part by determining the modularity of supermolecular functional units. To uncover these factors, we assigned the SC bands as quantitative traits, with all bands quantified as "on" (+1) or "off" (0). These traits were then mapped for QTLs. For bands 4 and 5, the data indicated that they are driven by a locus on chromosome 17 (Fig. 6A), containing 35 genes. Notably, this region includes *Cox7a2l*, and overlaps with its cis-pQTLs (Fig. 5C). *Cox7a2l* has been recently indicated as causal for certain types of SC formation between different inbred mouse lines (*37*), although this effect has been debated (*38*). We thus sought to examine how this locus can affect specific SC formation and whether *Cox7a2l* is indeed causal.

In mammals, SCs are formed by different stoichiometric combinations of three of the five individual complexes in the electron transport chain—complexes I, III, and IV—although it is poorly understood how they are formed or how different complexes influence overall mitochondrial function (*39*). To determine the stoichiometry of the observed bands, we performed in-gel activity assays for complex I (CI), CIV, CI+IV, and CIII using eight strains: four with the B6 allele of COX7A2L (e.g., BXD39) and four with the D2 allele (e.g., BXD32) (Fig. 6, B and C). We observed that SC formations with multiple copies of CIV—bands 4 ($I+III_2+IV_2$) and 5 ($I+III_2+IV_3$)—are inhibited in strains with B6-type *Cox7a2l*. Furthermore, a large increase in free/unconjugated CIII (blue arrow) and CIV (orange arrow) was observed in those strains with B6-type *Cox7a2l*, indicating indeed a lack of the assembly of these complexes into SCs. This analysis also indicated other variable SC formations at lower molecular weights, particularly complex III_2+IV_1 (dashed red line and pink arrow, Fig. 6B). Western blot analysis for COX7A2L shows its presence in SC bands 4 and 5 in D2-type strains (along with band III_2+IV_1; this complex was "hidden" under CV in the total BN-PAGE) and its complete absence in B6-type strains (Fig. 6C). This effect of COX7A2L on SC formation is noted broadly across strains with B6 or D2 alleles of the gene (e.g., Fig. 5F). However, the variant COX7A2L isoform does not seem to influence the formation of SC bands 2 ($I+III_2$) or 3 ($I+III_2+IV_1$).

We next examined the possibility that the *Cox7a2l* may not be the driving factor for SC variation in the BXDs and that the nearby leucine-rich pentatricopeptide (PPR) motif–containing gene (*Lrpprc*) may be causal (*38*). *Lrpprc* is 1.2 Mb downstream from *Cox7a2l*, and four BXD strains examined have recombinations between this interval: BXD56 has the D2 allele of *Cox7a2l* and the B6 allele of *Lrpprc*, whereas BXD44, BXD49, and BXD99 have the opposite. For these strains, SC bands 4 and 5 are absent in BXD44, 49, and 99 and present in BXD56 (Fig. 5F), as expected if *Cox7a2l* is causal. Furthermore, neither the transcript nor protein measurements of *Lrpprc* yield QTLs in the BXD livers. This finding does not preclude the possibility that LRPPRC is involved in SC formation. However, it is not the causal gene for the variable SC patterns observed in the BXDs. To further investigate the effects of COX7A2L on SC formation, we extracted mitochondria from the hearts of the same individuals as the liver. Again, SC patterns were strikingly different depending on genotype (Fig. 6D), yet the SC bands representing III_2+IV_1, $I+III_2+IV_2$, and $I+III_2+IV_3$ are present in the hearts of strains with the B6 isoform of *Cox7a2l*, albeit at diminished levels compared with their D2-type counterparts. Taken together, these data provide substantial evidence to show that COX7A2L is involved in the formation of many CIV-containing SC formations, yet that its influence varies between tissues. Additionally, these data provide a conceptual advancement in the current knowledge of SC formations in B6 by showing that $I+III_2+IV_1$ is in fact present [previously reported as absent (*37*)] and that variants in COX7A2L are causal for many of the differences between B6 and other common inbred strains, particularly D2.

Discussion

We have examined genetically and environmentally variant cohorts of the murine BXD GRP to determine how changes in the genome and environment interact to influence cellular processes and overall variation in complex metabolic traits. To model the molecular factors underlying phenotypic differences across the BXD population, we have applied an in-depth, multilayered approach including genetic, transcriptomic, proteomic, and metabolomic measurements. Systems-level technologies now permit the multilayered measurements of thousands of molecules associated with many physiological processes, at high throughput and with a high degree of quantitative accuracy and reproducibility. We show the first application of SWATH-MS in a diverse mammalian population by quantifying 2622 proteins measured in all 80 cohorts. As in earlier smaller-scale studies (40), genes' transcript levels are only moderate predictors for the protein levels. The predictive value depends strongly on how variable the transcript (or protein) is. Studies that induce massive transcriptional changes with gain- or loss-of-function techniques can rely on the fact that the resulting mRNA change will almost invariably be reflected in the corresponding protein's level. In contrast, it cannot be assumed that relatively subtle expression changes in a particular transcript will manifest at its protein level. This latter situation is particularly critical for in vivo population studies, where the top leads identified through microarray or RNA sequencing (RNA-seq) analyses frequently have far more modest differences than findings from in vitro studies. Likewise, genetic variants driving differential transcript expression (e.g., cis-eQTLs) are furthermore only infrequently mirrored at the protein level (e.g., matched cis-pQTLs) and vice versa. Measurement of both transcriptomics and proteomics in tandem appears essential because each measurement level unveils different aspects of the cellular state and regulatory mechanisms.

This greater scope of data analysis allows the identification of hundreds of causal genetic factors that regulate individual transcript and protein levels (i.e., QTLs), as well as for metabolites and phenotypes. This protein analysis allowed novel identification of variants of ECI2 not predicted

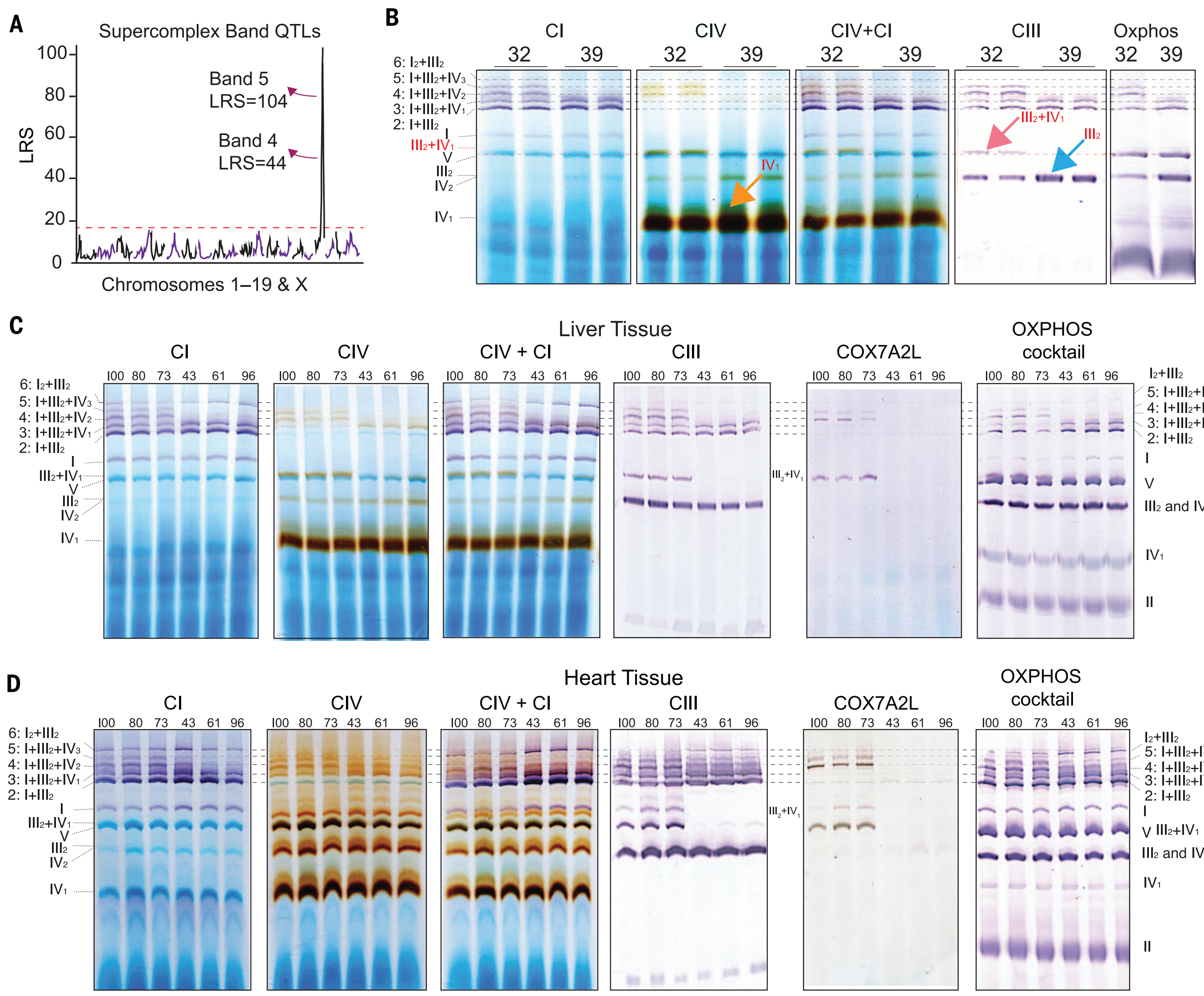


Fig. 6. Tissue variance in SC formation. (**A**) SC bands 4 and 5 mapped significantly as cQTLs to a locus on chromosome 17. (**B**) In-gel activity assays were performed in the liver tissues to determine SC's composition and relation to COX7A2L. Bands 2 to 5 could be identified confidently as CI + $CIII_2$ + variable numbers of CIV (0 to 3). (**C**) In-gel activity assays from livers of six additional BXD strains—three with the B6 allele of *Cox7a2l* (BXD73, BXD80, and BXD100) and three with the D2 allele (BXD43, BXD61, and BXD96). COX7A2L is present in bands 4 and 5 for strains with the D2 allele. (**D**) In-gel activity assays from hearts of the same individuals as above. COX7A2L is absent in bands 4 and 5 and III_2+IV_1 in strains with the B6 allele and present in strains with the D2 allele. Unlike liver, bands 4 and 5 are observed in all strains, albeit at lower levels in strains with the B6 allele of *Cox7a2l*, indicating tissue-specific differences in SC formation.

by genome or transcript data, as well as further delineation of variants affecting the expression of the four proteins in the BCKDC—effects not visible at the transcript level—which lead to a mild form of MSUD in the BXDs. In another, we could readily identify the causal factors driving variance in the metabolite D2HG to the protein D2HGDH. Moreover, however, the increased scope of these data facilitates the modeling and analysis of entire pathways. The PPAR and cholesterol biosynthesis pathways are highly variable in the BXDs due to both genetic and environmental factors and are known to influence the development of metabolic diseases, including fatty liver. Furthermore, we were able to use network analysis to identify *Mmab* and *Echdc1* as likely cholesterol-related genes, which we confirmed through in vitro analysis. For the oxphos gene network, the BXDs display strong levels of variation in both gene expression and the overall mitochondrial assembly of complexes in the ETC. Using the proteomic data, we identified COX7A2L as causal of major variants in SC organization—particularly, a lack of three specific SC bands (III_2+IV_2, I+III_2+IV_2, and I+III_2+IV_3)—and a consequent increase of the unconjugated levels of complexes IV_1, III_2, and IV_2. However, the mechanism of assembly appears strongly tissue dependent: These SCs can be formed in the heart, even in the absence of COX7A2L. Notably, the patterns of mitochondrial complexes are consistent across biological replicates, indicating that the many differences across strains and tissue are due primarily to differential regulation of mitochondria by the nuclear genome. In each of these pathways, the proteomic and metabolomic data extend gene-phenotype links that were previously identified at the transcript level but that were incomplete. To move forward in the analysis of mitochondria and associated disorders, it is hence necessary to analyze the protein levels of all regulators, as well as genetic, environment, and tissue-specific variants. Such implementations of new omics layers will not supersede the now-standard genomic and transcriptomic data sets. Rather, a combined transomic approach can fill in blind spots and assist in defining more detailed metabolic pathways.

Methods

Population handling

BXD strains were sourced from the University of Tennessee Health Science Center (Memphis, TN, USA) and bred at the École Polytechnique Fédérale de Lausanne (EPFL) animal facility for more than two generations before incorporation into the study. We examined 80 cohorts of the BXD population from 41 strains—41 on CD, 39 on HFD—with male mice from each strain separated into two groups of about five mice for each diet (two strains on HFD were lost before tissue collection). We started with 201 CD and 185 HFD mice, and a total of 183 CD and 168 HFD mice survived until they were killed at 29 weeks of age, with all cohorts having three or more individuals surviving to the end except BXD56 HFD, which had two. Strains were entered into the phenotyping program randomly and had staggered entry, typically by 2 weeks. Most strains entered with both dietary cohorts at the same time, with the exception of BXD50, 68, 69, 71, 84, 85, 89, 95, 96, and 101, where CD cohorts entered before HFD cohorts. All cohorts consisted of littermates. HFD feeding started at 8 weeks of age. Cohorts were communally housed by strain and diet from birth until 23 weeks of age and were then individually housed until they were killed at 29 weeks of age. CD is Harlan 2018 (6% kCal of fat, 20% kCal of protein, and 74% kCal of carbohydrates), and HFD is Harlan 06414 (60% kCal of fat, 20% kCal of protein, and 20% kCal of carbohydrates). All mice were housed under 12 hours of light alternated with 12 hours of dark, with ad libitum access to food and water at all times, except before they were killed, when mice were fasted overnight. All mice were housed in isolator cages with individual air filtration, except during the activity wheel test (10 days) when mice were in open-air cages in a room reserved for that test, after which, mice were returned to the filtered isolator cages. Body weight was measured weekly from 8 weeks of age until killing. Killing took place from 9:00 a.m. until 10:30 a.m., with isoflurane anesthesia followed by a complete blood draw (~1 mL) from the vena cava, followed by perfusion with phosphate-buffered saline. Half of the blood was placed into lithium-heparin (LiHep)–coated tubes and the other half in EDTA-coated tubes; then both were shaken and stored on ice, followed immediately by collection of the liver. The LiHep blood taken for plasma analysis was also centrifuged at 4500 revolutions per minute (rpm) for 10 min at 4°C before being flash-frozen in liquid nitrogen. Whole blood taken for cellular analysis was processed immediately after the killing (i.e., after ~1 to 2 hours on ice). Gallbladders were removed, and the livers were cut into small pieces before freezing in liquid nitrogen until preparation into mRNA, protein, or metabolite samples. Liver and blood serum were then stored at –80°C until analysis. All research was approved by the Swiss cantonal veterinary authorities of Vaud under licenses 2257 and 2257.1.

Phenotyping experiments

A visual summary of the phenotyping program is also included in Fig. 1B. At 16 weeks of age, after 8 weeks of dietary treatment, the cohorts underwent their first phenotyping test: 48 hours of respiration measurements in individual metabolic cages (Oxymax/CLAMS, Columbus Instruments). The first 24 hours were considered adaptation, and the second 24 hours were used for data analysis, including analysis of movement, the volume of oxygen inhaled, the volume of carbon dioxide exhaled, and derived parameters of these two, such as the respiratory exchange ratio (RER). One week later, all cohorts underwent an oral glucose tolerance test. Mice were fasted overnight before the test, and fasted glucose was tested with a glucometer at the tail vein. All individuals were then weighed and given an oral gavage of 20% glucose solution at 10 mL per kg of weight. Glucometer strips were used at 15, 30, 45, 60, 90, 120, 150, and 180 min after the gavage to examine glucose response over time. Blood was also collected at 0 (pregavage), 15, and 30 min to examine insulin levels. Two weeks later, at 19 weeks of age, we performed a noninvasive blood pressure measurement using a tail-cuff system (BP-2000 Blood Pressure Analysis System, Series II, Visitech Systems) over 4 days. The first 2 days were considered as adaptation to the apparatus, and the second 2 days were used for data analysis, and all measurements (systolic blood pressure, diastolic blood pressure, and heart rate) were averaged across both days. Outliers on a per-measurement basis were removed, but outlier mice were retained. Two weeks later, at 21 weeks of age, we performed a cold response test. The basal body temperatures of mice were examined rectally, after which mice were placed individually in prechilled cages in a room at 4°C. The cages were the standard housing cages but with only simple woodchip bedding, without supplement (e.g., tissue paper). Body temperature was checked every hour for 6 hours, after which the mice were returned to their normal housing cage. Two weeks later, at 23 weeks of age, the mice were placed individually in regular housing cages for basal activity recording. The housing cages were then placed in laser detection grids developed by TSE Systems (Bad Homburg, Germany). Within the cages, woodchip bedding was retained, but tissue bedding was removed (as it interferes with the laser detection). Food and water were as normal throughout the standard housing, both of which require rearing to reach. The detection grid has two layers: one for detecting X-Y movement ("ambulation") the other for Z movement ("rearing"). Both measurements are technically independent, although the measurements of movement are strongly correlated (r ~ 0.70). Mice were housed individually for the 48-hour experiment starting at about 10 a.m., with the night cycles (7 p.m. to 7 a.m., with 30 min of both dawn and dusk) used for movement calculations. We have recently published more interpretation and examination of the results from this experiment (*17*). After this 2-day experiment, all mice performed a VO_{2max} treadmill experiment using the Metabolic Modular Treadmill (Columbus Instruments). For the first 15 min in the machine for each individual, the treadmill was off while basal respiratory parameters were calculated. The last 2 minutes of data before the treadmill turned on are considered basal levels (most mice spend the first few minutes exploring the device). The treadmill then started at a pace of 4.8 m per minute (m/min), followed by a gradual increase over 60 s to 9 m/min, then 4 min at that pace before increasing to 12 m/min over 60 s, then four min at that pace before increasing to 15 m/min over 60 s, then 4 min at that pace, then the speed increased continuously by 0.015 m per second (or +0.9 m/min) thereafter until the end of the experiment at 63.5 min, 1354.5 m, or when the mouse is exhausted. CD cohorts ran against a 10° incline, whereas HFD cohorts were set at 0°. For this test, no mice reached the maximum distance recorded by the machine—all were taken out when exhausted. The distance, maximum VO_2, and

maximum RER were recorded. Maximums must be consistent across multiple measurements, and not single-measurement spikes, which were removed. Immediately after the treadmill experiment, mice were placed in individual open-air cages with ad libitum access to activity running wheels (Bioseb BIO-ACTIVW-M, Vitrolles, France). The final 24 hours of activity wheel access were recorded for all strains. For certain strains, all 10 days of activity wheel usage was recorded (depending on the availability of the recording system). After the 10th day, at ~25 weeks of age, mice underwent an identical treadmill experiment to that described above at 23 weeks of age. At this point, with the 10 days of voluntary training, three mice "completed" the experiment—two DBA/2Js on HFD and one BXD81 on CD. As before, the test was stopped for all other mice when they had reached exhaustion (considered as falling off the treadmill and inability to recover and continue running). After this experiment, mice were returned to their standard housing cages—individually—for 4 weeks. Mice were fasted overnight before they were killed. Details about killing are described in the previous section. In addition to the body weight measurements taken each week and before each phenotyping experiment, body composition was recorded at 16, 23, and 25 weeks of age—before respiration and the two VO_{2max} experiments. To do so, each mouse was placed briefly in an EchoMRI (magnetic resonance imaging) machine (the 3-in-1, Echo Medical Systems), where lean and fat mass are recorded, along with total body weight, taking ~1 min per individual. Lean mass is used as a corrective factor for respiratory calculations from the Comprehensive Lab Animal Monitoring System (CLAMS). All other tests are normalized to total body weight in our analyses.

Genomics

The parental lines of C57BL/6J and DBA/2J have been previously sequenced (13). Earlier genotype data—~8000 single-nucleotide polymorphisms (SNPs) per line—have been published previously (42). We have made use of a newer build of the genotype, using ~500,000 SNPs per line (unpublished), which helped refine recombination breakpoints, such that ~99.99% of the genotype of all BXD strains could be inferred. Full sequence data on the parental lines was published separately (18). The lower density (3806 markers) is available on GeneNetwork as well: www.GeneNetwork.org/genotypes/BXD.geno.

BXD sample preparation and analysis

For mRNA, 100-mg pieces of liver tissue were suspended in TRIzol (Invitrogen) and homogenized with stainless steel beads using a TissueLyser II (Qiagen) at 30 Hz for 2 min, followed by a standard phase separation extraction using chloroform and precipitated by isopropanol. mRNA concentration was measured for all samples and then pooled equally for each cohort (i.e., five biological replicates for BXD103 CD became one mixed pool of BXD103 CD). Pooled RNA was cleaned up using RNEasy (Qiagen). The mRNA of all cohorts was prepared in direct series over a ~2-week period. Seventy-six of the 77 cohorts had high-quality mRNA based on RNA integrity numbers ≥ 8.0, indicating that they are suitable for amplification and subsequent microarray analysis. Arrays were run for all cohorts in direct series over a 3-week period using the Affymetrix MouseGene 1.0 ST array at the Molecular Resource Center of Excellence in the University of Tennessee Health Science Center. Data were normalized using the robust multiarray average method (43), then analyzed in GeneNetwork and R.

For liver protein, the ~100-mg liver sample was homogenized with 4-mL radioimmune precipitation assay-modified buffer (1% Nonidet P-40, 0.1% sodium deoxycholate, 150 mM NaCl, 1 mM EDTA, 50 mM tris, pH 7.5, protease inhibitors EDTA-free, 10 mM NaF, 10 mM sodium pyrophosphate, 5 mM 2-glycerophosphate) in a glass-glass tight Dounce homogenizer (Wheaton Science Products) at 4°C. After the homogenates were centrifuged (20,000 g at 4°C for 15 min), the supernatant was collected and kept at 4°C. The pellets were resuspended with urea-tris buffer (50 mM tris, pH 8.1, 75 mM NaCl, 8 M urea, EDTA-free protease inhibitors, 10 mM NaF, 10 mM sodium pyrophosphate, 5 mM 2-glycerophosphate) and sonicated for 5 min, then centrifuged at 20,000 g for 15 min at 4°C. The supernatants from the two steps were combined, and protein concentrations were determined with the bicinchoninic acid protein assay (Thermo Fisher Scientific). For the precipitation and digestion of proteins in each sample, 200 μg of protein was precipitated with six volumes of ice-cooled acetone and kept 16 hours at –20°C. Then proteins were resuspended in 8 M urea/0.1 M NH_4HCO_3 buffer, reduced with 12 mM dithiotreitol for 30 min at 37°C, then alkylated with 40 mM iodoacetamide for 45 min at 25°C, in the dark. Samples were diluted with 0.1-M NH_4HCO_3 to a final concentration of 1.5-M urea, and sequencing grade porcine trypsin (Promega) was added to a final enzyme:substrate ratio of 1:100 and incubated for 16 hours at 37°C. Peptide mixtures were cleaned by Sep-Pak tC18 cartridges (Waters, Milford, MA, USA) and eluted with 40% acetonitrile. The resulting peptide samples were evaporated on a vacuum centrifuge until dry, then resolubilized in 2% acetonitrile/0.1% formic acid to 1μg/μL concentration.

For liver metabolites, the ~100-mg liver pieces were homogenized in 1 mL of 70% ethanol at –20°C. Metabolites were extracted by adding 7 mL of 70% ethanol at 75°C for 2 min. Extracts were centrifuged for 10 min at 4000 rpm at 4°C. Clean metabolite extracts were dried in a vacuum centrifuge and resuspended in double-distilled H_2O, with volume according to the weight of the extracted liver piece. Quantification of metabolites was performed on an Agilent 6550 quadrupole orthogonal acceleration–time-of-flight (Q-TOF) instrument by flow injection analysis time-of-flight mass spectrometry (24). All samples were injected in duplicates. Ions were annotated based on their accurate mass and the Human Metabolome Database reference list (44), allowing a tolerance of 0.001 Da. Unknown ions and those annotated as adducts were discarded. Theoretical m/z ratios—beyond the significant digits from the measurement sensitivity—are used as the unique index in the data files and online on GeneNetwork. For example, deprotonated fumarate corresponds to 115.0036897_MZ, malate to 133.0142794_MZ, α-ketoglutarate to 145.0141831_MZ, and D2HG to 147.0298102_MZ.

For blood serum analysis, samples were frozen in liquid nitrogen until large "batches" were ready, which were run in multiples of 16 samples. Samples were then thawed on ice, diluted 1:1 in NaCl solution, and then processed on a Dade Behring Dimension Xpand Clinical Chemistry System. Sixteen metabolites were measured based on standard reagent-reaction spectrophotometry. Due to the long period of time for this study, two chemical batch effects were noted for HDL, free fatty acids, aspartate transaminase, lactate dehydrogenase, and creatinine measurements. These metabolite measurements separated distinctly into two batches based on the time of measurement and a change in the batch of reagent used. To account for this, the two batches for these five metabolites were *Z*-score normalized and then combined, losing information about absolute values but retaining utility for correlation analyses.

BN-PAGE and in-gel activity

For SC analysis, mitochondria were isolated, protein was extracted, and these extracts were prepared and run on BN-PAGE, described in detail in a separate methods paper (45). In brief, ~30 mg of tissue was homogenized and taken for mitochondrial isolation. For BN-PAGE, 50 or 35 μg of mitochondria from liver and heart, respectively, was solubilized in digitonin and sample buffer (Invitrogen, BN 2008). For the liver, these samples were the same tissues used for omics analysis, using all CD cohorts with three biological replicates per cohort. For the heart, these were the same mice as for the liver. Digitonin/protein ratio of 4 g/g was used for liver and 8 g/g was used for heart (for better band resolution, because heart contains more SCs than liver). Electrophoresis was performed using Native PAGE Novex Bis-Tris Gel System (3 to 12%), as per manufacturer's instructions with minor modification. Gel transfer was performed using Invitrogen iBlot gel transfer system. For detection of the complexes, anti-oxphos cocktail (Invitrogen, 457999) and WesternBreeze Chromogenic Western Blot Immunodetection Kit (Invitogen, WB7103) were used. In the final detection step, incubation of the membrane with the chromogenic substrate was for 8 min for all the gels. Membranes were dried, scanned, and each visible SC band was independently scored from 1 to 5. All samples were then run across several gels, and we observed nearly complete biological reproducibility (heritability) for band presence or absence. Contrast across gels varied significantly; thus, bands were categorized in a binary manner as "present" or "not present" for QTL analysis.

For in-gel activity assays, electrophoresis was performed for 3 hours (30 min at 150 V and 2.5 hours at 250 V). Complex I activity was performed by incubating the gels for 15 to 30 min in the substrate composed of 2 mM tris-HCl pH 7.4;

0.1 mg/mL NADH, and 2.5 mg/mL nitrotetrazolium blue. CIV activity was performed by incubating the gels for 30 to 40 min in the substrate composed of 25 mg of 3,3′-diamidobenzidine tetrahydrochloride; 50 mg cytochrome c; 45 mL of 50 mM phosphate buffer pH 7.4, and 5 mL water. CIV+CI activity was performed by subsequently incubating the gels in the substrate for CIV followed by incubation in CI. All reactions were stopped with 10% acetic acid.

BN-PAGE was run as well for six cohorts in the HFD state. We observed no clear difference across diets, and no difference related to the COX7A2L-dependent bands. In-gel activity assays were run for CI, CIV, CIV+CI, and CIII for eight strains (four with the B6 allele of *Cox7a2l* and four for the D2 allele).

Proteomics: Peptide library development

To develop peptide libraries, we chose 58 cohorts and used 100 μg of protein lysate each (digested as described above). The resulting peptides were mixed and loaded for off-gel electrophoresis fractionation as previously described (*46*). The 24 fractions were combined into 10 fractions and cleaned up with C18 column. Each fraction was analyzed with classical shotgun data acquisition with a AB Sciex TripleTOF 5600 mass spectrometer interfaced to an Eksigent NanoLC Ultra 2D Plus high-performance liquid chromatography system. Samples were loaded on to a PicoFrit emitter coupled with an analytical column (75 μm diameter) with buffer A (2% acetonitrile, 0.1% formic acid) and eluted with a 135-min linear gradient of 2 to 35% buffer B (90% acetonitrile, 0.1% formic acid) with a flow rate of 300 nL/min. The 20 most intense precursors with charge states 2 to 5 were selected for fragmentation, and the MS2 spectra were acquired in the range 50 to 2000 m/z for 100 ms, and precursor ions were excluded from reselection for 15 s.

Profile mode wiff files from shotgun data acquisition were transformed to centroid mode and converted to mzML files using AB Sciex Data Converter, and then converted to mzXML files using FileConverter. The mzXML files were searched against the canonical UniProt complete proteome database for mouse using the Trans-Proteomic Pipeline (*47*). A decoy database was generated by reversing the amino acid sequences and appended to the target database. Cysteine carboxy-amidomethylation was set as the static modification, and methionine oxidation was set as the variable modification. Peptides with up to one missed cleavage site were allowed. Mass tolerance was set to 25 parts per million for precursor ions and 0.4 Da for fragment ions. The pepXML files were combined using iProphet (*48*), and the integrated pepXML file was used to generate the redundant spectra library containing all peptide spectra matches using SpectraST (*49*). Retention time of peptide identification was transformed to indexed retention time (iRT) values based on the linear regression calibrated for each shotgun run using the information of the spiked iRT peptides. The median of iRT values of each peptide were calculated using in-house script, and the consensus library was constructed using SpectraST. We then selected the top five most abundant b and y fragment ions of each peptide to generate the assays for SWATH-MS targeted extraction. The target assay library contains the precursor and fragment m/z values and the relative intensities of the fragment ions, as well as the average iRT value of each precursor. Decoy assays were appended to the target assay library for error rate estimation.

Proteomics: SWATH mass spectrometry and targeted data extraction

SWATH-MS represents the next generation in large-scale quantitative proteomics measurement techniques and provides a substantial leap in both scope and quality over the most commonly used untargeted proteomics technique today, discovery proteomics (also known as shotgun proteomics). Although discovery proteomics achieves high proteome coverage, the identification and quantification are biased toward those proteins with higher abundance in the sample, and it suffers from inherently poor reproducibility when large sample cohorts are being analyzed. This hurdle has limited the implementation of this approach in large population studies. Recently, targeted proteomics methods have been developed to increase the reproducibility of proteome measurement, such as selected reaction monitoring [used in our previous study (*19*)]. Due to lower throughput, however, studies using these alternative techniques inevitably measure fewer proteins than studies using shotgun proteomics. Recently, we have developed SWATH, which has demonstrated the ability to quantify thousands of proteins with good reproducibility and quantification accuracy across large sample cohorts (*3*). Consequently, SWATH provides considerable improvements in both proteome coverage and measurement reproducibility.

The SWATH-MS was performed with the 5600 TripleTOF mass spectrometer, as previously described (*3*). The chromatographic parameters were as described above. For SWATH-MS–based experiments, the mass spectrometer was operated in a looped product ion mode and specifically tuned to allow a quadrupole resolution of 25 Da/mass selection. Using an isolation width of 26 Da (containing 1 Da for the window overlap), a set of 32 overlapping windows was consecutively constructed covering the 400 to 1200 m/z precursor range (*3*). The collision energy (CE) for each window was determined based on appropriate collision energy for a charge 2+ ion centered upon the window with a spread of 15 eV. An accumulation time (dwell time) of 100 ms was used for all fragment ion scans in high-sensitivity mode. The survey scans were acquired in high-resolution mode at the beginning of each SWATH-MS cycle, resulting in a duty cycle of 3.4 s.

The SWATH-MS results were first converted to profile mzXML files using ProteoWizard (*50*). The SWATH-MS targeted data extraction was performed using OpenSWATH workflow (*51*), which applies a target-decoy scoring model to estimate the false discovery rate (FDR) by the mProphet algorithm (*52*). Retention time alignment between SWATH maps was performed based on the clustering of reference peptides using a nonlinear alignment algorithm (*53*). Fragment ion chromatograms were extracted according to the target-decoy assay library with a width of 0.05 m/z, and peak groups were scored based on the elution profile of fragment ions, similarity of elution time and relative intensities with the assay library, and the properties of the tandem MS spectrum extracted at the chromatographic peak apex. Finally, peptide FDR was estimated according to the score distribution of target and decoy assays.

Proteomics: SWATH protein classification and quality control

A key in all protein measurement techniques, including antibody-based approaches, is that they must choose and quantify a specific, small subset of the protein's overall amino acid sequence to analyze. ProteinProphet analysis on the data from OpenSWATH ensures that the peptides identified are proteotypic (*54*). The majority of the resulting quantified peptides—20,718 of 22,208—are uniquely attributable to a single protein. The remaining 1510 match common regions of up to nine distinct proteins; these peptides were discarded from analysis in this study. All 22,208 peptides are recorded and available in the raw file download on GeneNetwork.org. Peptide quantities were calculated with imsbInfer, an R package (https://github.com/wolski/imsbInfer). We furthermore analyzed several peptide sequences known to target amino acid sequences with missense mutations in the BXD (e.g., the peptide SAVYPTSAVQMEAALR for the gene *Mrsa* has a M to L variant at the highlighted amino acid). In this case, we observed striking differences in the alleles that were not observed for other peptides measuring the same gene that do not have the missense mutation. Thus, 100% amino acid matches are necessary for reliable detection, indicating that our unique peptides were accurately assigned. These missense mutations can lead to false-positive cis-pQTLs; thus, in cases where only one peptide mapped to a cis-pQTL, we controlled for sequence variants, often highlighting missense mutations.

After performing all of these controls, we were typically left with multiple peptide measurements corresponding to a single protein (a bit over seven peptides per protein on average, although this number is highly variable). To reduce multiple testing in subsequent multilayered analyses, we thus assigned the "best" peptide to represent the overall gene level. This was done through several sequential criteria: (i) peptides mapping as cis-pQTLs, (ii) peptides with at least nominally significant correlation to transcript ($P < 0.05$), and (iii) peptides correlating with known controls in independent layers of data [e.g., HMGCS1 should fluctuate with mevalonate (*55*) (Fig. 2E)]. A total of 632 proteins were assigned based on criteria i, and a further 824 were assigned based on criteria ii, whereas only a single protein was assigned based on criteria iii. The remaining 1165 proteins were assigned to peptides based on intensity—a standard selection criterion in MS analyses (*56*).

cis-pQTL and mRNA-peptide correlations were performed at a per-peptide level. Due to the increased multiple testing issue of taking all individual peptides, network analyses were performed with only one peptide per gene (i.e., per protein), as were trans-pQTL analyses and all other transomic analyses. It is worth noting that this obstacle of multiple peptides per protein is analogous to pioneering quantitative polymerase chain reaction experiments and microarray design, where there were also considerable difficulties and trial-and-error methods for choosing the "correct" primers (or probe sets or RNA-seq reads) that most correctly measure overall transcript levels (*57*). Although nearly two decades of transcriptome research has led to the fairly reliable establishment of which RNA fragments correspond properly to the overall transcript, no similar database is yet available for proteins. However, as we generally observe strong networks of associated genes (e.g., Fig. 4), we can determine that the peptides, by and large, accurately represent the protein levels. In time, with many studies such as this one, we expect that it will become feasible to determine guidelines and databases for how to best determine overall protein level from individual peptides.

Cholesterol validation

Huh7 and HepG2 cells were grown in Dulbecco's minimum essential medium (DMEM) or MEM, respectively, and supplemented with 10% fetal bovine serum. Cells were treated for 48 hours (drugs) or 72 hours (siRNA knockdown) before harvesting and preparation of peptides for mass spectrometric analysis. Control conditions for drug experiments consist of untreated and 0.1% dimethyl sulfoxide–treated cells; control conditions for siRNA-knockdown experiments consist of cells that were untreated, mock treated (only lipofectamine RNAiMAX transfection reagent), or cells treated with negative-control siRNA (antisense strand: CUACGAUAGACCGGUCGUAtt). Silencer Select siRNAs were used in a concentration of 5 nM. Knockdown of proteins was performed for LDLR (low-density lipoprotein receptor) with siRNAs s224006 and s224007; for SREBF2 with siRNA s27 and s28; for SREBF (sterol regulatory element binding transcription factor) with a mix of s27, s28, s129, and s130, targeting both SREBF1 and SREBF2. For the LPDS + statin condition, cells were incubated in medium containing 10% LPDS and 1 or 5 μM atorvastatin. Protein signal was quantified for three peptides for MMAB and ECHDC1 and five peptides for HMGCS1, and signal was normalized to control conditions. Experiments were performed for all conditions in two or three biological replicates for siRNA or drug experiments, respectively. The box plots consist, therefore, of two (SREBF si), four (SREBF2 si, LDLR si), or six (drug controls, siRNA controls, and LPDS statin) data points.

Metabolomics in liver: Metabolite classification and quality control

We identified 979 unique metabolite features based on m/z using flow-injection TOF-MS. Of these features, 699 could be attributed to a single metabolite, including in cases where of the two "possible" enantiomers, one is more predominant than the other (e.g., L versus D amino acids). The remaining 280 metabolites were "clusters" with no clear predominant feature—for example, the "glucose" metabolite measurements could not be separated from fructose, galactose, or mannose measurements, as all share the same m/z. The "main" metabolite and all possible alternatives are listed with the data on GeneNetwork.org for the raw file download (press the "INFO" button next to the data set on the main search page and download the data set and supplemental data files).

Normality and significance

Outliers were Windsorized for QTL mapping, which was performed using R/qtl (*58*). The normality of data was checked by the Shapiro-Wilk test in R, with $W \geq 0.90$ considered normal. Student's t test was used for two-group comparisons in normal data, and Welch's t test was used otherwise. Bonferroni's correction was used for multiple testing for tests of correlation. FDR calculations were used for peptide scoring. cis-QTLs used a LRS = 12 as the significance threshold, whereas trans-QTLs used a LRS = 18, as cis-QTLs do not need to correct for multiple locus testing across the genome. Raw P values and corrected P values are both reported when applied.

Figure preparation

Network graphs (Figs. 1I; 4, A, C, and G; and 5A) were performed using Spearman correlation, keeping all edges with P values less than or equal to the reported cutoff in the panel legend. These panels were made in R using the custom package imsbInfer, currently on Github (https://github.com/wolski/imsbInfer) but in the process of being added to Bioconductor. The Circos plot (Fig. 5B) was generated using Circos (http://www.circos.ca). Spearman correlation matrices (Figs. 3B and 4F) were generated in R using the corrgram package. Metabolite structures were drawn with ChemBioDraw (Fig. 3G). QTL plots were drawn in R/qtl (e.g., Fig. 3, A, D, and I). Most other figures (e.g., Fig. 1, C to H and J) were generated using standard R plotting packages included in gplots or ggplot2—e.g., stripchart, plotCI, and barplot2. Final figures were all prepared with Adobe Illustrator. Western blots and BN-PAGE gels were transferred and scanned, then cropped for the figure preparation without other rearrangement or editing, except for Fig. 5F, where biological triplets were "cut" and rearranged based on numerical order of the BXDs. For this panel, contrast was edited individually to obtain an even tone across the full complement of strains. Because this is across several independent gels, the relative intensity across strains is unreliable, hence the binary assignment of SC presence rather than quantity.

REFERENCES AND NOTES

1. E. S. Lander, Initial impact of the sequencing of the human genome. *Nature* **470**, 187–197 (2011). doi: 10.1038/nature09792; pmid: 21307931
2. D. C. Sévin, A. Kuehne, N. Zamboni, U. Sauer, Biological insights through nontargeted metabolomics. *Curr. Opin. Biotechnol.* **34**, 1–8 (2015). doi: 10.1016/j.copbio.2014.10.001; pmid: 25461505
3. L. C. Gillet *et al*., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **11**, O111.016717 (2012). doi: 10.1074/mcp.O111.016717; pmid: 22261725
4. N. Selevsek *et al*., Reproducible and consistent quantification of the *Saccharomyces cerevisiae* proteome by SWATH-mass spectrometry. *Mol. Cell. Proteomics* **14**, 739–749 (2015). doi: 10.1074/mcp.M113.035550; pmid: 25561506
5. S. F. Kingsmore, I. E. Lindquist, J. Mudge, W. D. Beavis, Genome-wide association studies: Progress in identifying genetic biomarkers in common, complex diseases. *Biomark. Insights* **2**, 283–292 (2007). pmid: 19662211
6. M. Mayr, B. Madhu, Q. Xu, Proteomics and metabolomics combined in cardiovascular research. *Trends Cardiovasc. Med.* **17**, 43–48 (2007). doi: 10.1016/j.tcm.2006.11.004; pmid: 17292045
7. A. R. Joyce, B. O. Palsson, The model organism as a system: Integrating 'omics' data sets. *Nat. Rev. Mol. Cell Biol.* **7**, 198–210 (2006). doi: 10.1038/nrm1857; pmid: 16496022
8. E. A. Franzosa *et al*., Sequencing and beyond: Integrating molecular 'omics' for microbial community profiling. *Nat. Rev. Microbiol.* **13**, 360–372 (2015). doi: 10.1038/nrmicro3451; pmid: 25915636
9. M. R. Emmett *et al*., Integrative biological analysis for neuropsychopharmacology. *Neuropsychopharmacology* **39**, 5–23 (2014). doi: 10.1038/npp.2013.156; pmid: 23800968
10. R. Romero *et al*., The use of high-dimensional biology (genomics, transcriptomics, proteomics, and metabolomics) to understand the preterm parturition syndrome. *BJOG* **113** (suppl. 3), 118–135 (2006). doi: 10.1111/j.1471-0528.2006.01150.x; pmid: 17206980
11. X. Wang *et al*., Joint mouse-human phenome-wide association to test gene function and disease risk. *Nat. Commun.* **7**, 10464 (2016). doi: 10.1038/ncomms10464; pmid: 26833085
12. D. M. Altshuler *et al*., Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010). doi: 10.1038/nature09298; pmid: 20811451
13. N. Wong, G. Morahan, M. Stathopoulos, J. Proietto, S. Andrikopoulos, A novel mechanism regulating insulin secretion involving Herpud1 in mice. *Diabetologia* **56**, 1569–1576 (2013). doi: 10.1007/s00125-013-2908-y; pmid: 23620059
14. P. A. Andreux *et al*., Systems genetics of metabolism: The use of the BXD murine reference panel for multiscalar integration of traits. *Cell* **150**, 1287–1299 (2012). doi: 10.1016/j.cell.2012.08.012; pmid: 22939713
15. I. E. Lush, The genetics of tasting in mice. VI. Saccharin, acesulfame, dulcin and sucrose. *Genet. Res.* **53**, 95–99 (1989). doi: 10.1017/S0016672300027968; pmid: 2744455
16. T. P. Dalton *et al*., Identification of mouse SLC39A8 as the transporter responsible for cadmium-induced toxicity in the testis. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3401–3406 (2005). doi: 10.1073/pnas.0406085102; pmid: 15722412
17. E. G. Williams *et al*., An evolutionarily conserved role for the aryl hydrocarbon receptor in the regulation of movement. *PLOS Genet.* **10**, e1004673 (2014). doi: 10.1371/journal.pgen.1004673; pmid: 25255223
18. T. M. Keane *et al*., Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289–294 (2011). doi: 10.1038/nature10413; pmid: 21921910
19. Y. Wu *et al*., Multilayered genetics and omics dissection of mitochondrial activity in a mouse reference population. *Cell* **158**, 1415–1430 (2014). doi: 10.1016/j.cell.2014.07.039
20. M. S. Van Der Knaap *et al*., D-2-Hydroxyglutaric aciduria: Biochemical marker or clinical disease entity? *Ann. Neurol.* **45**, 111–119 (1999). doi: 10.1002/1531-8249(199901)45:1<111::AID-ART17>3.0.CO;2-N; pmid: 9894884
21. S. C. Hasenfuss *et al*., Regulation of steatohepatitis and PPARγ signaling by distinct AP-1 dimers. *Cell Metab.* **19**, 84–95 (2014). doi: 10.1016/j.cmet.2013.11.018; pmid: 24411941
22. M. Inoue *et al*., Increased expression of PPARgamma in high fat diet-induced liver steatosis in mice. *Biochem. Biophys. Res. Commun.* **336**, 215–222 (2005). doi: 10.1016/j.bbrc.2005.08.070; pmid: 16125673
23. A. Ghazalpour *et al*., Comparative analysis of proteome and transcriptome variation in mouse. *PLOS Genet.* **7**, e1001393 (2011). doi: 10.1371/journal.pgen.1001393; pmid: 21695224
24. T. Fuhrer, D. Heer, B. Begemann, N. Zamboni, High-throughput, accurate mass metabolome profiling of cellular extracts by

flow injection-time-of-flight mass spectrometry. *Anal. Chem.* **83**, 7074–7080 (2011). doi: 10.1021/ac201267k; pmid: 21830798

25. J. Li-Hawkins *et al*., Cholic acid mediates negative feedback regulation of bile acid synthesis in mice. *J. Clin. Invest.* **110**, 1191–1200 (2002). doi: 10.1172/JCI0216309; pmid: 12393855
26. K. A. Strauss, E. G. Puffenberger, D. H. Morton, in *GeneReviews*, R. A. Pagon *et al*., Eds. (Seattle, WA, 2013), http://www.ncbi.nlm.nih.gov/books/NBK1319.
27. C. B. Newgard *et al*., A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* **9**, 311–326 (2009). doi: 10.1016/j.cmet.2009.02.002; pmid: 19356713
28. D. L. Aylor, Not just another trait: Methods for the genetic analysis of gene expression (ProQuest Information and Learning Company, North Carolina State University, 2008); http://www.lib.ncsu.edu/resolver/1840.16/3804.
29. Y. Achouri *et al*., Identification of a dehydrogenase acting on D-2-hydroxyglutarate. *Biochem. J.* **381**, 35–42 (2004). doi: 10.1042/BJ20031933; pmid: 15070399
30. M. Kanehisa *et al*., Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Res.* **42**, D199–D205 (2014). doi: 10.1093/nar/gkt1076; pmid: 24214961
31. L. Izem *et al*., Effect of reduced low-density lipoprotein receptor level on HepG2 cell cholesterol metabolism. *Biochem. J.* **329**, 81–89 (1998). doi: 10.1042/bj3290081; pmid: 9405278
32. H. Scharnagl *et al*., Effect of atorvastatin, simvastatin, and lovastatin on the metabolism of cholesterol and triacylglycerides in HepG2 cells. *Biochem. Pharmacol.* **62**, 1545–1555 (2001). doi: 10.1016/S0006-2952(01)00790-0; pmid: 11728391
33. J. L. Goldstein, M. S. Brown, A century of cholesterol and coronaries: From plaques to genes to statins. *Cell* **161**, 161–172 (2015). doi: 10.1016/j.cell.2015.01.036; pmid: 25815993
34. T. M. Teslovich *et al*., Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010). doi: 10.1038/nature09270; pmid: 20686565
35. C. J. Willer *et al*., Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274–1283 (2013). doi: 10.1038/ng.2797; pmid: 24097068
36. M. Bonke *et al*., Transcriptional networks controlling the cell cycle. *Genes Genomes Genetics* **3**, 75–90 (2013). doi: 10.1534/g3.112.004283; pmid: 23316440
37. E. Lapuente-Brun *et al*., Supercomplex assembly determines electron flux in the mitochondrial electron transport chain. *Science* **340**, 1567–1570 (2013). doi: 10.1126/science.1230381; pmid: 23812712
38. A. Mourier, S. Matic, B. Ruzzenente, N. G. Larsson, D. Milenkovic, The respiratory chain supercomplex organization is independent of COX7a2l isoforms. *Cell Metab.* **20**, 1069–1075 (2014). doi: 10.1016/j.cmet.2014.11.005; pmid: 25470551
39. E. A. Schon, N. A. Dencher, Heavy breathing: Energy conversion by mitochondrial respiratory supercomplexes. *Cell Metab.* **9**, 1–3 (2009). doi: 10.1016/j.cmet.2008.12.011; pmid: 19117538
40. S. P. Gygi, Y. Rochon, B. R. Franza, R. Aebersold, Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730 (1999). doi: 10.1128/MCB.19.3.1720; pmid: 10022859
41. J. A. Vizcaíno *et al*., ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223–226 (2014). doi: 10.1038/nbt.2839; pmid: 24727771
42. J. L. Peirce, L. Lu, J. Gu, L. M. Silver, R. W. Williams, A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genet.* **5**, 7 (2004). doi: 10.1186/1471-2156-5-7; pmid: 15117419
43. R. A. Irizarry *et al*., Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003). doi: 10.1093/biostatistics/4.2.249; pmid: 12925520
44. D. S. Wishart *et al*., HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res.* **41**, D801–D807 (2013). doi: 10.1093/nar/gks1065; pmid: 23161693
45. P. Jha, X. Wang, J. Auwerx, Analysis of mitochondrial respiratory chain supercomplexes using blue native polyacrylamide gel electrophoresis (BN-PAGE). *Curr. Protoc. Mouse Biol.* **6**, 1–14 (2016). doi: 10.1002/9780470942390.mo150182; pmid: 26928661
46. P. Picotti, B. Bodenmiller, L. N. Mueller, B. Domon, R. Aebersold, Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell* **138**, 795–806 (2009). doi: 10.1016/j.cell.2009.05.051; pmid: 19664813
47. A. Keller, J. Eng, N. Zhang, X. J. Li, R. Aebersold, A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol. Syst. Biol.* **1**, 2005.0017 (2005). doi: 10.1038/msb4100024; pmid: 16729052
48. D. Shteynberg *et al*., iProphet: Multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell. Proteomics* **10**, M111.007690 (2011). doi: 10.1074/mcp.M111.007690; pmid: 21876204
49. H. Lam *et al*., Development and validation of a spectral library searching method for peptide identification from MS/MS. *Proteomics* **7**, 655–667 (2007). doi: 10.1002/pmic.200600625; pmid: 17295354
50. D. Kessner, M. Chambers, R. Burke, D. Agus, P. Mallick, ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536 (2008). doi: 10.1093/bioinformatics/btn323; pmid: 18606607
51. H. L. Röst *et al*., OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat. Biotechnol.* **32**, 219–223 (2014). doi: 10.1038/nbt.2841; pmid: 24727770
52. L. Reiter *et al*., mProphet: Automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* **8**, 430–435 (2011). doi: 10.1038/nmeth.1584; pmid: 21423193
53. H. Weisser *et al*., An automated pipeline for high-throughput label-free quantitative proteomics. *J. Proteome Res.* **12**, 1628–1644 (2013). doi: 10.1021/pr300992u; pmid: 23391308
54. A. I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658 (2003). doi: 10.1021/ac0341261; pmid: 14632076
55. M. S. Brown, J. L. Goldstein, The SREBP pathway: Regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**, 331–340 (1997). doi: 10.1016/S0092-8674(00)80213-5; pmid: 9150132
56. J. E. Elias, F. D. Gibbons, O. D. King, F. P. Roth, S. P. Gygi, Intensity-based protein identification by machine learning from a library of tandem mass spectra. *Nat. Biotechnol.* **22**, 214–219 (2004). doi: 10.1038/nbt930; pmid: 14730315
57. M. A. Stalteri, A. P. Harrison, Interpretation of multiple probe sets mapping to the same gene in Affymetrix GeneChips. *BMC Bioinformatics* **8**, 13 (2007). doi: 10.1186/1471-2105-8-13; pmid: 17224057
58. D. Arends, P. Prins, R. C. Jansen, K. W. Broman, R/qtl: High-throughput multiple QTL mapping. *Bioinformatics* **26**, 2990–2992 (2010). doi: 10.1093/bioinformatics/btq565; pmid: 20966004

ACKNOWLEDGMENTS

We thank S. Lamy, C. Cartoni, and the Center of Phenogenomics (CPG, EPFL) for their help on the phenotyping and L. Rose at the Molecular Resource Center of Excellence at the University of Tennessee Health Science Center for processing the microarrays. J.A. is the Nestlé Chair in Energy Metabolism. Research was further supported by the EPFL, ETHZ, ERC (AdG-670821 and Proteomics v3.0; AdG-233226), the AgingX program of the Swiss Initiative for Systems Biology (51RTP0-151019, 2013/134, and 2013/153), SNSF (31003A-140780, 31003A-143914, and CSRII3-136201), and the NIH (R01AG043930). Raw transcriptome data are published on GEO: GSE60149. Raw MS data for SWATH proteomics are available on PRIDE (*41*): PXD003266. Raw MS data for metabolomics are available on MassIVE: MSV000079411. Genotype information is available at www.GeneNetwork.org/genotypes/BXD.geno. Full sequence data on the parental lines were published separately (*18*). All data from this study can also be found and queried on the www.GeneNetwork.org resource. On the main search page, set Species to "Mouse" and Group to "BXD." Change "Type" to "Phenotype" and search for LISP3 to find phenotype data. Transcriptome data can be downloaded on GeneNetwork by changing Type to "Liver mRNA" and Data Set to any of the four EPFL/LISP entries. Pressing the INFO button to the right will lead to a page with links to download the raw data sets in the upper right. The same process can be performed to download the metabolome (change Type to "Liver Metabolome") or proteome data (change Type to "Liver Proteome"). The semiprocessed data files used for each omics layer can also be found in the supplementary materials online.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/aad0189/suppl/DC1
Data Files 1 to 8

15 July 2015; accepted 15 April 2016
10.1126/science.aad0189

REPORTS

GEOPHYSICS

Deeper penetration of large earthquakes on seismically quiescent faults

Junle Jiang[1]* and Nadia Lapusta[1,2]†

Why many major strike-slip faults known to have had large earthquakes are silent in the interseismic period is a long-standing enigma. One would expect small earthquakes to occur at least at the bottom of the seismogenic zone, where deeper aseismic deformation concentrates loading. We suggest that the absence of such concentrated microseismicity indicates deep rupture past the seismogenic zone in previous large earthquakes. We support this conclusion with numerical simulations of fault behavior and observations of recent major events. Our modeling implies that the 1857 Fort Tejon earthquake on the San Andreas Fault in Southern California penetrated below the seismogenic zone by at least 3 to 5 kilometers. Our findings suggest that such deeper ruptures may occur on other major fault segments, potentially increasing the associated seismic hazard.

The style of faulting in Earth's crust is widely accepted to be depth-dependent, with an upper layer that produces earthquakes and a lower layer that predominantly deforms stably (*1*). The upper layer is commonly referred to as the "seismogenic zone," and the geodetically estimated boundary between the two is commonly called the "locking depth," because the seismogenic zone is often locked in the interseismic period. The faulting transition with depth is dominated by temperature and is due to bulk properties transitioning from purely elastic to inelastic and to quasi-static fault friction properties transitioning from velocity-weakening to velocity-strengthening (Fig. 1). Major strike-slip faults feature extreme localization of slip at seismogenic depths (*2*), as well as continuing localization of the deformation even below the seismogenic zone, based on studies of deep tectonic tremor (*3*), postseismic deformation (*4*), and exhumed faults (*5*); we refer to this deeper localization as "deeper creeping fault extensions."

During the quasi-static interseismic periods between major earthquakes, these deeper creeping fault extensions should continuously load the adjacent locked fault areas and induce microseismicity there, due to the typical assumption that locked areas are seismogenic. The microseismicity at depth is indeed observed on some fault segments, most notably the Parkfield segment of the San Andreas Fault (SAF) (*6*) (Fig. 2). Such pronounced and concentrated microseismicity that occurs persistently over time should be commonly observed on faults. Yet several stretches of the SAF, including the Cholame, Carrizo, Mojave, and Coachella segments, are seismically quiescent (devoid of small earthquakes) in their interseismic periods, with negligible seismic moment release as compared to the active segments of the fault (*7*) (Fig. 3). The quiescence over most of the seismogenic zone for such mature faults can be due to their low stress in comparison to their static strength. However, the fault areas right next to the deeper creeping fault extensions should be well stressed and produce microseismicity regardless of whether the shallower fault regions are quiescent or not.

Here we show that the absence of concentrated microseismicity at the bottom of the seismogenic zone on mature fault segments can be due to the deeper penetration of (previous) large earthquakes. We have conjectured this relation based on the following rather general mechanical consideration. If the locked-creeping transition and the associated concentrated and continuous stressing are at the boundary of, or within, the seismogenic zone capable of nucleating seismic events, then one will expect the concentrated stressing to cause microseismicity even on homogeneous faults (*8*), but especially in the presence of heterogeneity of fault properties and stresses. However, if dynamic earthquake rupture penetrates below the seismogenic zone, it could drop stress in the ruptured creeping areas, making them effectively locked and placing the locked-creeping transition at a depth below the seismogenic zone, where the associated concentrated stressing is unlikely to initiate seismic events. As a result, fault segments with deeper slip in large events would lack microseismicity at greater depths, at least until the locked-creeping transition, which would become shallower with time due to reloading by deeper creep, reaches the seismogenic zone. This argument holds regardless of whether the deeper creeping fault extensions are governed by frictional slip (as explored

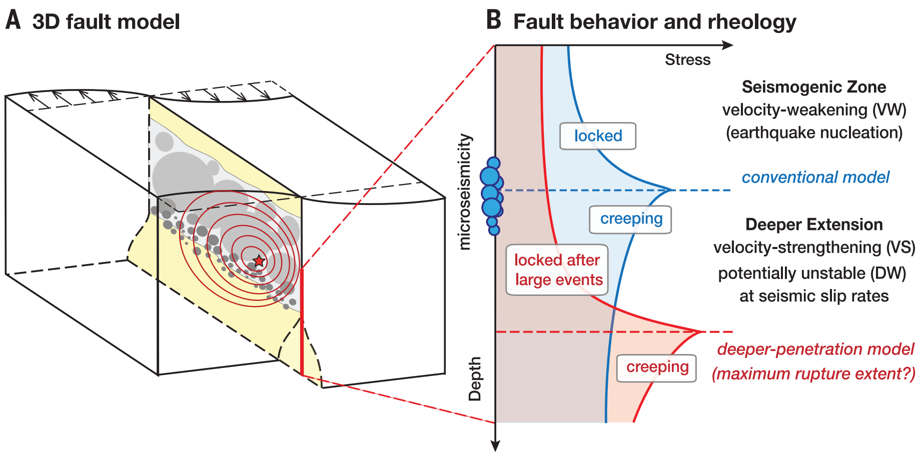


Fig. 1. Schematic illustration of our fault model and the locked-creeping transition. (**A**) A strike-slip fault model with the seismogenic zone (light gray areas), creeping regions (yellow), and fault heterogeneity (dark gray circles). The initiation point and rupture fronts of a large earthquake are illustrated by the red star and contours, respectively. (**B**) The locked seismogenic zone and creeping regions below are typically interpreted as having VW and VS rate-and-state friction properties, respectively. In purely rate-and-state models, the VW/VS boundary and locked-creeping transition nearly coincide, and the associated concentrated shear stressing induced at the locked-creeping transition (blue line) promotes microseismicity at the bottom of the seismogenic zone in the interseismic period (blue circles). However, large earthquake rupture may extend seismic slip deeper than the VW/VS boundary, due to enhanced dynamic weakening (DW) at high slip rates, putting the locked-creeping transition and the associated concentrated stressing (red line) within the VS region and hence suppressing microseismicity nucleation.

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA. [2]Division of Engineering and Applied Sciences, California Institute of Technology, Pasadena, CA, USA.
*Present address: Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. **†Corresponding author. Email: lapusta@caltech.edu**

in this work) or inelastic (e.g., viscoelastic or plastic) flow. As long as the fault extensions are sufficiently localized right below the seismogenic zone, as supported by multiple lines of evidence (*3–5*), the loading they impose on the seismogenic zone and its consequences should be the same.

This insight sheds light on the depth extent of large earthquakes, which is important for understanding deep crustal faulting, earthquake scaling relations, and fault segment interactions but is still poorly constrained. Inversions for fault-slip distribution of recent large strike-slip earthquakes usually do not provide reliable constraints on the depth extent of coseismic slip, due to their overall non-uniqueness as well as the decrease of imaging resolution with depth (*9*, *10*). Moreover, monitoring of fault segments with high seismic hazard, such as the SAF in California, is often limited to the late interseismic periods of large events (*11*). Meanwhile, a growing number of studies have been challenging the notion that dynamic slip during earthquakes is always confined within the seismogenic zone. Geological field studies report the overprinting of natural pseudotachylytes on mylonitic zones, attributed to repeated seismic slip overlapping with aseismic creep below the seismogenic zone (*12*), in accordance with the transitional regime with semi-brittle deformation mechanisms in conceptual fault models based on experimental and field studies (*1*, *13*). Deeper penetration of larger earthquakes can also explain the observed slip-length scaling of large events (*14*).

Dynamic rupture propagation into the deeper creeping zones is possible, based on our current laboratory-based understanding of fault friction, which has been gaining acceptance and validation through the comparison of earthquake models with observations. At low slip rates of 10^{-9} to 10^{-2} m/s, consistent with plate motion, earthquake nucleation, and postseismic slip, friction has been successfully described by logarithmic rate-and-state friction laws (*10*, *15*). Such laws interpret the seismogenic zones as areas of velocity-weakening (VW) properties that allow for earthquake nucleation, and the other fault areas as having velocity-strengthening (VS) properties that promote stable creep (Fig. 1). Models with the rate-and-state friction reproduce a wide range of fault behaviors, including earthquake

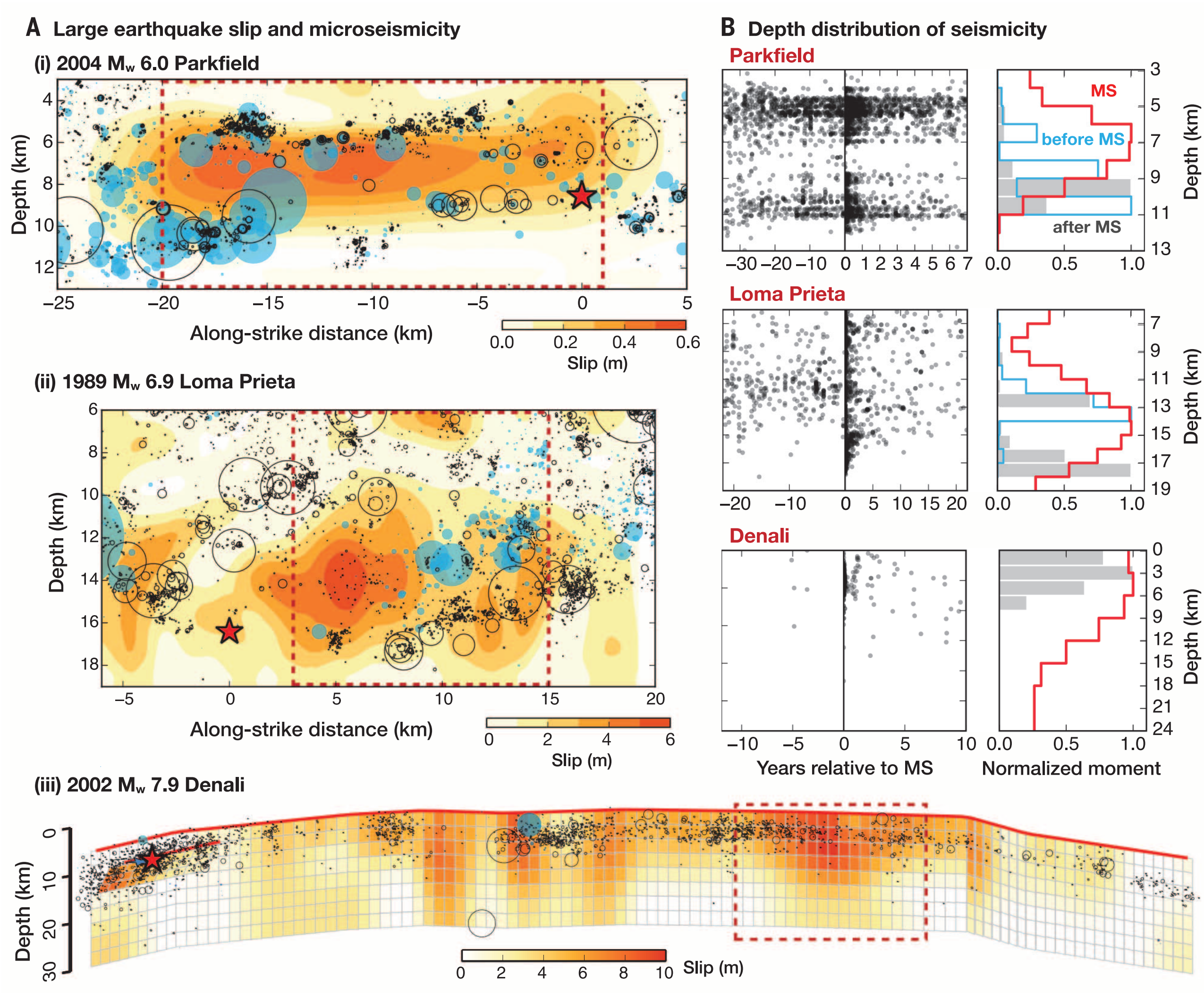


Fig. 2. Observations of large earthquakes and microseismicity patterns on major strike-slip faults. (**A**) Spatial relations of the inferred coseismic slip during large earthquakes (in color, with hypocenters as red stars) and microseismicity before (blue circles) and after (black circles), over time periods shown in (B). The large earthquakes are: (i) 2004 M_w 6.0 Parkfield (*6*, *16*), (ii) 1989 M_w 6.9 Loma Prieta (*32*), and (iii) 2002 M_w 7.9 Denali (*33*). Small earthquakes within 2, 4, and 5 km of the fault for the three cases, respectively, are projected onto the fault plane (except iii) and plotted using a circular crack model with the same seismic moment and 3 MPa stress drop. (**B**) (Left) Time evolution of the depths of seismicity (gray circles) and (right) the depth distribution of normalized total seismic moment released before (blue lines), during (red lines), and after (gray) the mainshock (MS). We considered seismicity and coseismic fault slip inside the regions of largest slip outlined by the red dashed lines in (A). Seismic moment release before the Denali event is not shown because of the small number of events.

sequences and aseismic slip (*16*). However, at slip rates of ~10^{-1} m/s and higher, enhanced dynamic weakening of fault frictional resistance, amply documented in high-velocity laboratory experiments (*17*) and supported by theoretical studies (*18*), could dominate earthquake rupture propagation. When an earthquake reaches deeper fault extensions, increased strain rate and shear heating could lead to strain localization and dynamic weakening (*19*), effectively turning the creeping fault regions into seismic ones (*20*).

We confirmed the hypothesized relation between the depth of coseismic slip in large earthquakes and microseismicity patterns, by numerical simulations of earthquake sequences in two fault models with the laboratory-derived friction laws (Fig. 4). In model M1, dynamic weakening is restricted to occur within the VW region, resulting in earthquake rupture confined within the seismogenic zone, whereas model M2 has dynamic weakening extended deeper into the VS regions below, allowing deeper earthquake rupture. We used the thermal pressurization of pore fluids (*18*, *21*) as the dynamic weakening mechanism, because fluids can be present at deeper fault extensions; however, the qualitative results of the models should be similar for other dynamic weakening mechanisms. The depth extent of efficient dynamic weakening due to thermal pressurization of pore fluids depends on a number of factors, including the shearing zone width (*5*), permeability, the extent of the inelastic dilatancy (*22*), and the effectiveness of pore pressure in reducing the effective normal stress (*23*). In both models, fault heterogeneity that could generate microseismicity is represented by VW patches with nucleation sizes smaller than that of the larger-scale VW region. Although the fault heterogeneity is likely to be more complex, we use the patches and put them only around the VW/VS transition for numerical efficiency. Our simulations are quite challenging, as they reproduce all stages of earthquake sequences, including spontaneous earthquake nucleation, dynamic rupture propagation with full inclusion of wave-mediated stress effects, and aseismic slip (*16*, *20*). We describe the numerical methods and model parameters in the supplementary materials (*10*).

The two models demonstrate the conjectured relation between the depth of coseismic slip in large earthquakes, microseismicity patterns, and the locked-creeping transition (Fig. 4, B to D). The transition is defined here as the fault depth with slip rates of 10% of $V_{\max}$, the maximum slip rate over the fault at the time, which has the physical significance of approximately corresponding to the depth of the highest concentrated stressing. This definition is different from the conventional locking depth inverted from surface geodetic observations (*24*), which interprets the actual depth distribution of slip rates in terms of a simplified dislocation model with a fully locked shallower layer and a fully creeping deeper fault extension. Coseismic slip of large events penetrates into the deeper fault extensions in model M2, but the coseismic slip is largely confined within the seismogenic zone in model M1, as intended. Correspondingly, in M1, the locked-creeping transition is at the bottom of the seismogenic zone immediately after the large event, causing abundant microseismicity throughout the interseismic period. In M2, however, the locked-creeping transition is below the seismogenic zone during most of the interseismic period, leading to a small number of interseismic events on the VW patches positioned below the large-scale VW/VS transition. The locked-creeping transition migrates up-dip over time, and the migration can be approximately predicted based on the earthquake stress drop $\Delta\tau$, product $(a-b)\overline{\sigma}$ of the VS friction properties and effective normal stress, fault recurrence interval, and long-term fault-slip rate (*10*).

Observed microseismicity patterns before and after major earthquakes on tectonic faults further support our hypothesis (Fig. 2) (*10*). On some fault segments, concentrated microseismicity occurs at what appear to be rheological transitions, with an increased activity below the ruptured region of, and after, a major event, such as the 2004 moment magnitude (M_w) 6.0 Parkfield and 1984 M_w 6.2 Morgan Hills earthquakes (*10*). In such cases, the slip in the major event probably occurs above the deeper concentrated microseismicity. For larger events, such as the 1989 M_w 6.9 Loma Prieta earthquake, one also observes the occurrence of microseismicity at depth before the mainshock and increased activity after the event, with some variability in local fault areas. In sharp contrast with these smaller events, microseismicity at depth is largely absent before or after all recent major (M_w > 7.5) strike-slip earthquakes that we have considered, including the 2002 M_w 7.9 Denali and 1999 M_w 7.6 Izmit earthquakes (*10*). According to our models, the absence of microseismicity means that these earthquakes ruptured into the creeping fault extensions, which is more likely for larger events with larger slip. Larger slip at depth and a larger depth extent of the rupture may promote larger slip on shallower fault areas as well, making the fault segment more prone to quiescence at all depths. At the same time, the difference in microseismicity patterns between faults with smaller and larger events is most evident for their deeper parts, whereas the microseismicity in the shallower fault regions is more variable, pointing to a stronger influence of other factors, such as variations of fault structure and properties. The lack of on-fault aftershocks after some of the large strike-slip events was previously attributed to the relatively uniform fault friction, as evidenced by supershear rupture propagation (*25*). This is consistent with our conclusions, because relatively uniform fault conditions, such as a

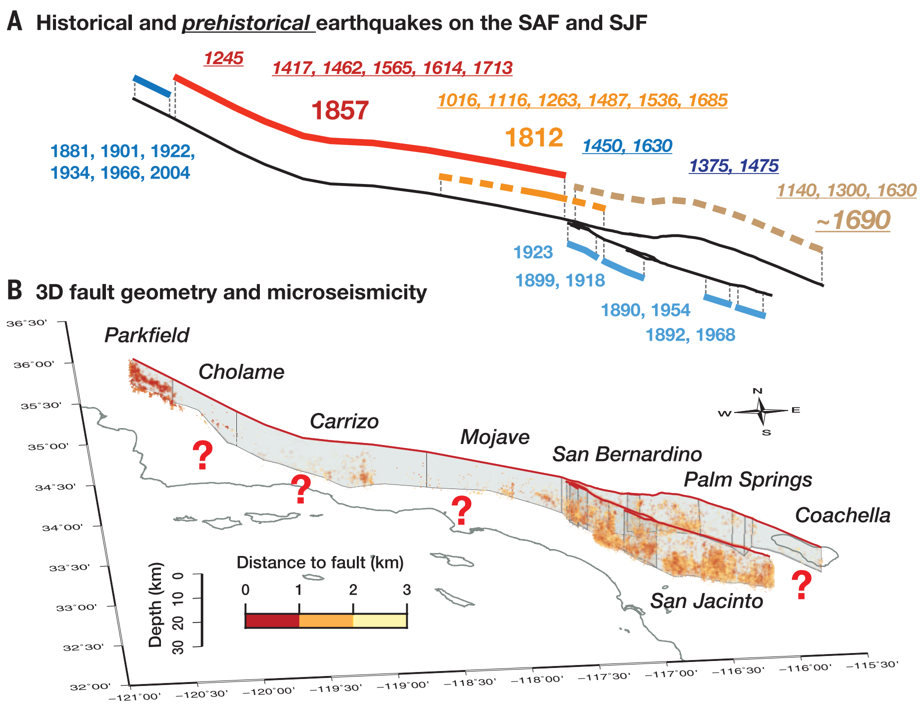


Fig. 3. Microseismicity and the potential for deeper ruptures on the SAF and the San Jacinto Fault (SJF) in Southern California. (**A**) Historical and prehistorical earthquakes on the SAF and SJF, with approximate rupture extent for major events (solid and dashed lines are for well-documented and uncertain cases, respectively) (*10*). Approximate calendar years for prehistorical events are only shown for the SAF in underlined italics. (**B**) Seismicity (1981–2011) within 3 km from the SAF and SJF (*6*, *7*). Active seismicity at depth is observed on the Parkfield and San Bernardino segments of the SAF and on the SJF. The Cholame, Carrizo, Mojave, and Coachella segments of the SAF have been seismically quiet for decades. The 1857 and ~1690 events probably penetrated below the seismogenic zone, and similar behavior can occur in future events.

smooth fault geometry, are likely to promote not only supershear transition but also enhanced dynamic weakening, larger slip, and hence its deeper penetration. More generally, our observations show the absence of aftershocks near regions of large deep slip, regardless of whether the rupture was supershear or not, suggesting that our model provides an additional explanation for the lack of aftershocks.

The relation between the microseismicity and depth of slip in large earthquakes helps us understand historical events and evaluate potential future earthquake scenarios on major strike-slip fault segments. The 1857 M_w 7.9 Fort Tejon earthquake is the last major event on the SAF/San Jacinto fault system in Southern California (*7*) (Fig. 3) that ruptured the Cholame, Carrizo, and Mojave segments (*26*, *27*). The last major earthquake on the Coachella segment occurred in ~1690 and potentially also ruptured the San Bernardino and Palm Springs segments (*11*). The recurrence of such events poses severe seismic hazards for Southern California. Virtually no microseismicity is currently observed on all these segments (Fig. 3). In light of our modeling, this observation implies that, ~150 to ~300 years after the previous major seismic events, the locked-creeping transition on those segments is still below the bottom of the seismogenic zone. To achieve that, dynamic rupture on those segments should have penetrated an additional depth below the seismogenic zone, at least 3 to 5 km based on our physical model (*10*) and perhaps much more. Interseismic geodetic observations indeed suggest that the Carrizo, Mojave, and Coachella segments are accumulating more potency deficit than other fault segments, which they are expected to release in future events (*28*). Deeper penetration of coseismic slip on the Cholame and Carrizo segments is consistent with the inference of much larger slip at depth, of approximately 11 and 16 m, respectively, than the 3- to 6-m slip at the surface during the 1857 event (*27*, *29*, *30*).

In summary, we suggest that the absence of microseismicity at the bottom of seismogenic zones points to a deeper rupture extent in recent major earthquakes, probably due to coseismic weakening of otherwise stable deeper regions. Furthermore, the deeper penetration may be quite common for large events on mature strike-slip

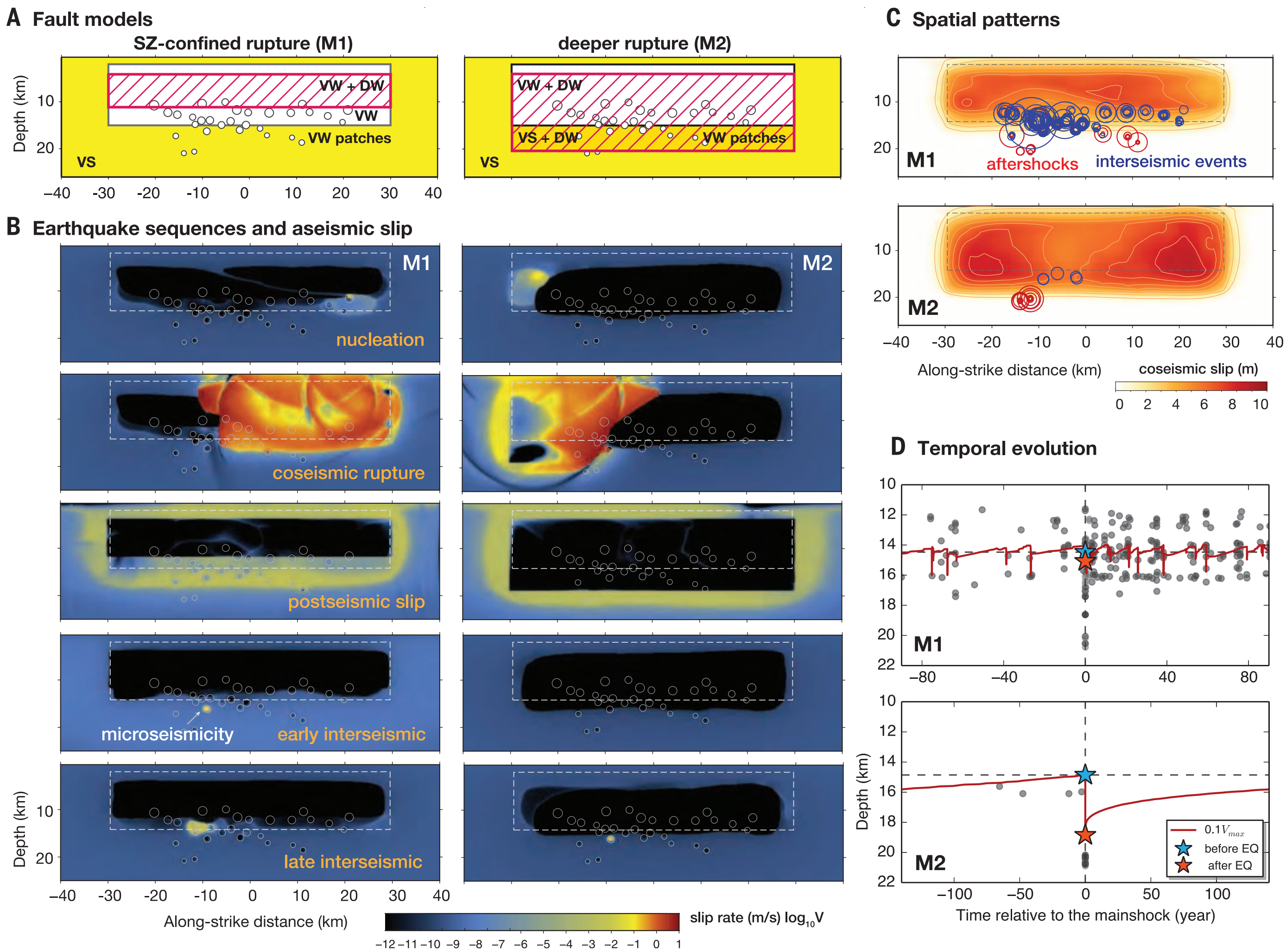


Fig. 4. The relation between the depth extent of large earthquakes and microseismicity in simulated earthquake sequences. (**A**) Model M1 has DW (red hashed region) within the VW region (white), with ruptures confined to the seismogenic zone (SZ). Model M2 has DW extending into the VS region (yellow), potentially allowing for deeper ruptures. VW circular patches of smaller nucleation sizes represent fault heterogeneity at the transitional depths. (**B**) Different stages in the long-term fault behavior illustrated by snapshots of fault-slip rates on a logarithmic scale. The two models differ in the coseismic rupture extent and the location of the locked-creeping transition with respect to the VW/VS boundary (white dashed outline), and hence in microseismicity activity. (**C**) Spatial patterns of microseismicity in the post- and interseismic periods of a typical large event (with coseismic slip in color), plotted using the same method as in Fig. 2. Note the concentrated microseismicity in M1 and its near-absence in M2. (**D**) Time evolution of the locked-creeping transition (red line) and seismicity depths (black dots). The blue and red stars represent the depth of the locked-creeping transition before and after the mainshock, respectively. The time windows equal the recurrence intervals (180 and 280 years, respectively). EQ refers to the large earthquake shown in (B).

faults. We have demonstrated this phenomenon in a friction-based fault model, but the overall dynamics of the process should be similar for viscoplastic deeper fault extensions, which may dynamically localize and weaken due to shear-heating and strain-rate effects during large earthquakes (*19*) and maintain their localization through the interseismic period because of the resulting structural differences in terms of their grain size and heterogeneity (*31*). Our study has focused on major strike-slip faults, but it has important relevance for the seismic hazard of megathrust subduction zones that are seismically quiescent, such as the Cascadia subduction zone, given the critical effect of down-dip rupture limit on coastal shaking.

REFERENCES AND NOTES

1. C. H. Scholz, *The Mechanics of Earthquakes and Faulting* (Cambridge Univ. Press, New York, ed. 2, 2002).
2. F. M. Chester, J. S. Chester, *Tectonophysics* **295**, 199–221 (1998).
3. D. R. Shelly, *Nature* **463**, 648–652 (2010).
4. R. E. Reilinger *et al.*, *Science* **289**, 1519–1524 (2000).
5. J. Cole *et al.*, *J. Geophys. Res. Solid Earth* **112**, B12304 (2007).
6. F. Waldhauser, W. L. Ellsworth, D. P. Schaff, A. Cole, *Geophys. Res. Lett.* **31**, L18608 (2004).
7. E. Hauksson, W. Yang, P. M. Shearer, *Bull. Seismol. Soc. Am.* **102**, 2239–2244 (2012).
8. N. Lapusta, J. R. Rice, *J. Geophys. Res. Solid Earth* **108**, 2205 (2003).
9. P. M. Mai, K. K. S. Thingbaijam, *Seismol. Res. Lett.* **85**, 1348–1357 (2014).
10. Supplementary materials are available on *Science* Online.
11. Y. Fialko, *Nature* **441**, 968–971 (2006).
12. A. Lin *et al.*, *Tectonophysics* **402**, 21–35 (2005).
13. T. Shimamoto, *J. Struct. Geol.* **11**, 51–64 (1989).
14. B. E. Shaw, S. G. Wesnousky, *Bull. Seismol. Soc. Am.* **98**, 1633–1641 (2008).
15. J. Dieterich, in *Treatise on Geophysics*, G. Schubert, Ed. (Elsevier, Oxford, ed. 2, 2007), vol. 4, pp. 93–110.
16. S. Barbot, N. Lapusta, J.-P. Avouac, *Science* **336**, 707–710 (2012).
17. G. Di Toro *et al.*, *Nature* **471**, 494–498 (2011).
18. J. R. Rice, *J. Geophys. Res.* **111**, B05311 (2006).
19. J. D. Platt, J. W. Rudnicki, J. R. Rice, *J. Geophys. Res. Solid Earth* **119**, 4334–4359 (2014).
20. H. Noda, N. Lapusta, *Nature* **493**, 518–521 (2013).
21. R. H. Sibson, *Nature* **243**, 66–68 (1973).
22. P. Segall, A. M. Bradley, *J. Appl. Mech.* **79**, 031013 (2012).
23. G. Hirth, N. M. Beeler, *Geology* **43**, 223–226 (2015).
24. J. C. Savage, *J. Geophys. Res.* **111**, B04405 (2006).
25. M. Bouchon, H. Karabulut, *Science* **320**, 1323–1325 (2008).
26. K. E. Sieh, *Bull. Seismol. Soc. Am.* **68**, 1421–1434 (1978).
27. O. Zielke, J. R. Arrowsmith, L. G. Ludwig, S. O. Akçiz, *Science* **327**, 1119–1122 (2010).
28. B. R. Smith-Konter, D. T. Sandwell, P. M. Shearer, *J. Geophys. Res.* **116**, B06401 (2011).
29. L. B. Grant, A. Donnellan, *Bull. Seismol. Soc. Am.* **84**, 241–246 (1994).
30. E. E. Runnerstrom, L. B. Grant, J. R. Arrowsmith, D. D. Rhodes, E. M. Stone, *Bull. Seismol. Soc. Am.* **92**, 2659–2669 (2002).
31. P. B. Kelemen, G. Hirth, *Nature* **446**, 787–790 (2007).
32. G. C. Beroza, *Bull. Seismol. Soc. Am.* **81**, 1603–1621 (1991).
33. J. L. Elliott, J. T. Freymueller, B. Rabus, *J. Geophys. Res.* **112**, B06421 (2007).

ACKNOWLEDGMENTS

This study was supported by the U.S. Geological Survey (USGS) (grant G14AP00033), the National Science Foundation (NSF) (grants EAR 1142183 and 1520907), and the Southern California Earthquake Center (SCEC, funded by NSF cooperative agreement EAR-0529922 and USGS Cooperative agreement 07HQAG0008). This is SCEC contribution no. 6139. Numerical simulations for this study were carried out on the CITerra Dell cluster at the Division of Geological and Planetary Sciences of the California Institute of Technology. We thank J.-P. Ampuero, J.-P. Avouac, E. Hauksson, and M. Simons for helpful discussions and comments on the manuscript. Earthquake catalogs and fault-slip models are compiled from published literature and publicly available sources. Numerical data are available from the authors upon request.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1293/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S21
Tables S1 to S3
Movies S1 and S2
References (*34–117*)

27 December 2015; accepted 11 May 2016
10.1126/science.aaf1496

QUANTUM SIMULATION

Quantum spin dynamics and entanglement generation with hundreds of trapped ions

Justin G. Bohnet,[1]* Brian C. Sawyer,[1,2] Joseph W. Britton,[1,3] Michael L. Wall,[4] Ana Maria Rey,[5] Michael Foss-Feig,[3,6] John J. Bollinger[1]*

Quantum simulation of spin models can provide insight into problems that are difficult or impossible to study with classical computers. Trapped ions are an established platform for quantum simulation, but only systems with fewer than 20 ions have demonstrated quantum correlations. We studied quantum spin dynamics arising from an engineered, homogeneous Ising interaction in a two-dimensional array of $^9Be^+$ ions in a Penning trap. We verified entanglement in spin-squeezed states of up to 219 ions, directly observing 4.0 ± 0.9 decibels of spectroscopic enhancement, and observed states with non-Gaussian statistics consistent with oversqueezed states. The good agreement with ab initio theory that includes interactions and decoherence lays the groundwork for simulations of the transverse-field Ising model with variable-range interactions, which are generally intractable with classical methods.

Quantum simulation, in which a well-controlled quantum system emulates another system to be studied, can be used to address classically intractable problems in fields including condensed-matter and high-energy physics, cosmology, and chemistry (*1–3*). Of particular interest are simulations of the transverse-field Ising spin model (*4*), described by the Hamiltonian

$$\hat{H}_{\mathrm{T}} = \hat{H}_{\mathrm{I}} + \hat{H}_{\mathrm{B}} \quad (1)$$

$$\hat{H}_{\mathrm{I}} = \frac{1}{N}\sum_{i<j}^{N} J_{i,j}\hat{\sigma}_i^z\hat{\sigma}_j^z,\ \hat{H}_{\mathrm{B}} = \sum_{i}^{N} B_x\hat{\sigma}_i^x \quad (2)$$

where N is the number of spins, $J_{i,j}$ parameterizes the spin-spin interaction, B_x parameterizes a transverse magnetic field, and $\hat{\sigma}^z$, $\hat{\sigma}^x$ are Pauli spin matrices. A quantum simulation of $\hat{H}_{\mathrm{T}}$ could illuminate complex phenomena in quantum magnetism, including quantum phase transitions, many-body localization, and glassiness in spin systems (*5–8*), and clarify whether quantum annealing can provide an advantage for solving hard optimization problems (*9*, *10*).

Ensembles of photons, ions, neutral atoms, molecules, and superconducting circuits are all being developed as quantum simulation platforms (*3*). A variety of quantum spin models have been realized with large ensembles of neutral atoms (*11–15*) and molecules (*16*). Trapped-ion quantum simulators can implement $\hat{H}_{\mathrm{T}}$ (*17–19*) and have a number of advantages over other implementations, such as high-fidelity state preparation and readout, long trapping and coherence times, and strong, variable-range spin-spin couplings. To date, trapped-ion simulators have been constrained to systems of about 20 spins (*18*, *20*), for which classical numerical simulations remain tractable; substantial engineering efforts are under way to increase the number of ions by cryogenically cooling linear traps and two-dimensional (2D) surface-electrode traps (*21*, *22*).

Penning traps have emerged as a viable option for performing quantum simulations with hundreds of ions (*23–26*). Laser-cooled ions in a Penning trap self-assemble into 2D triangular lattices and are amenable to similar high-fidelity spin-state control, long trapping times, and generation of transverse-field Ising interactions as ions in linear Paul traps. Previous work in Penning traps demonstrated control of the collective spin (*27*) and benchmarked the engineered, variable-range Ising interaction in the mean-field, semiclassical limit (*24–26*). However, for a simulator of quantum magnetism to be trusted, quantum

[1]National Institute of Standards and Technology (NIST), Boulder, CO 80305, USA. [2]Georgia Tech Research Institute, Atlanta, GA 30332, USA. [3]Army Research Lab, Adelphi, MD 20783, USA. [4]JILA, NIST, and University of Colorado, Boulder, CO 80309, USA. [5]JILA, NIST, and Department of Physics, University of Colorado, Boulder, CO 80309, USA. [6]Joint Quantum Institute and NIST, Gaithersburg, MD 20899, USA.
***Corresponding author. Email: justin.bohnet@nist.gov (J.G.B.); jjb@nist.gov (J.J.B.)**

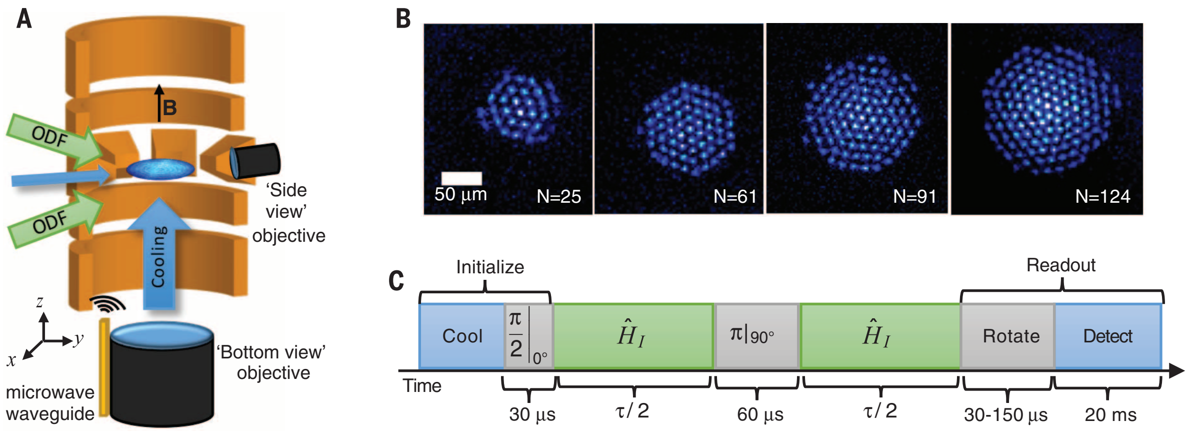


Fig. 1. Penning trap quantum simulator. (**A**) A cross-sectional illustration of the Penning trap (not to scale). The orange electrodes provide axial confinement and the rotating wall potential. The 4.5 T magnetic field is directed along the z axis. The blue disk indicates the 2D ion crystal. Resonant Doppler cooling is performed with the beams along z and y. The spin state–dependent optical dipole force (ODF) beams enter ±10° from the 2D ion plane. Resonant microwave radiation for coupling ground states $|\uparrow\rangle$ and $|\downarrow\rangle$ is delivered through a waveguide. State-dependent fluorescence is collected through the pair of imaging objectives, where the bright state corresponds to $|\uparrow\rangle$. (**B**) Coulomb crystal images in a frame rotating at ω_r with $^9Be^+$ ions in $|\uparrow\rangle$, with the number of ions N indicated. (**C**) The typical experiment pulse sequence, composed of cooling laser pulses (blue), microwave pulses (gray), and ODF laser pulses (green). Cooling and repumping initialize each ion in $|\uparrow\rangle$, and then a microwave $\pi/2$ pulse prepares the spins along the x axis. Suddenly switching on $\hat{H}_I$ initiates the non-equilibrium spin dynamics. The microwave π pulse implements a spin echo, reducing dephasing from magnetic field fluctuations and ODF laser light shifts. State readout consists of a final global rotation and fluorescence detection. The final microwave pulse area and phase are chosen to measure the desired spin projection.

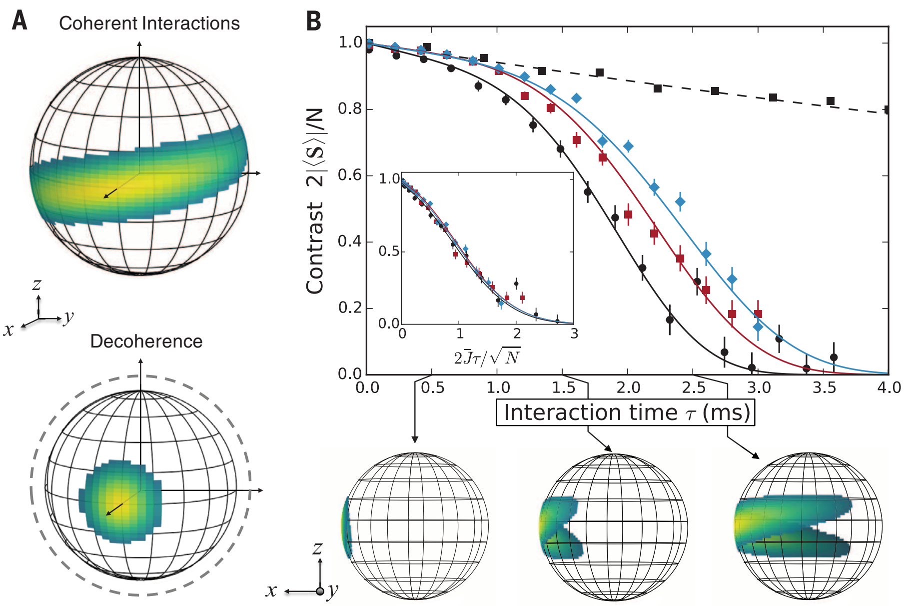


Fig. 2. Depolarization of the collective spin from spin-spin interactions and decoherence. (**A**) The Husimi distribution of the collective spin state on a Bloch sphere calculated for the experimental parameters in (B), with $N = 21$, illustrating (top) an oversqueezed state generated by the Ising interaction at time $\tau = 2$ ms with no decoherence and (bottom) a loss of contrast only from decoherence, effectively shrinking the Bloch sphere. (**B**) Contrast versus interaction time for $N = 21$, 58, and 144 ions indicated by black circles, red squares, and blue diamonds, respectively. Data are means ± SD; the solid lines are predictions, with no free parameters, from a model that includes decoherence from spontaneous emission (28). The contrast decay from decoherence caused by spontaneous emission is measured in the absence of spin-spin coupling (black squares with the dashed line showing an exponential fit). At each τ, the detuning δ is adjusted to eliminate spin-motion coupling at the end of the experiment, resulting in a different $\bar{J}\propto 1/\delta$ for each point. The Bloch spheres show the Husimi distribution for a pure state of $N = 21$ at three different interaction times, ignoring the effects of decoherence. Inset: The data collapse to a common curve with proper rescaling, indicating that the depolarization is dominated by coherent spin-spin interactions.

correlations generated by the Ising interaction must be observed and understood. For large trapped-ion simulators, this benchmarking requires a detailed accounting of many-body physics in an open quantum system.

Here, we observed and benchmarked entanglement in hundreds of trapped ions generated with engineered Ising interactions in a 2D array of $^9Be^+$ ions in a Penning trap. To enable efficient theoretical computation of the spin dynamics (28), we performed experiments with a homogeneous Ising interaction and without simultaneous application of the transverse field B_x, finding good agreement with a solution of the full quantum master equation.

Our experimental system consists of between 20 and 300 $^9Be^+$ ions confined to a single-plane Coulomb crystal in a Penning trap (Fig. 1) (28). The trap is characterized by an axial magnetic field $|\mathbf{B}| = 4.45$ T and an axial trap frequency $\omega_z = 2\pi \times 1.57$ MHz. A stack of cylindrical electrodes generates a harmonic confining potential along their axis. Radial confinement is provided by the Lorentz force from $\mathbf{E} \times \mathbf{B}$–induced rotation in the axial magnetic field. Time varying potentials applied to eight azimuthally segmented electrodes generate a rotating wall potential that controls the crystal rotation frequency ω_r, typically between $2\pi \times 172$ kHz and $2\pi \times 190$ kHz.

The spin-½ system is the $^2S_{1/2}$ ground state of the valence electron spin $|\uparrow\rangle \equiv |m_s\rangle = +^1/_2$, $|\downarrow\rangle \equiv |m_s\rangle = -^1/_2$. In the magnetic field of the Penning trap, the ground state is split by 124 GHz. A resonant microwave source provides an effective transverse field, which we use to perform global rotations of the spin ensemble with a Rabi frequency of 8.3 kHz. The T_2 spin echo coherence

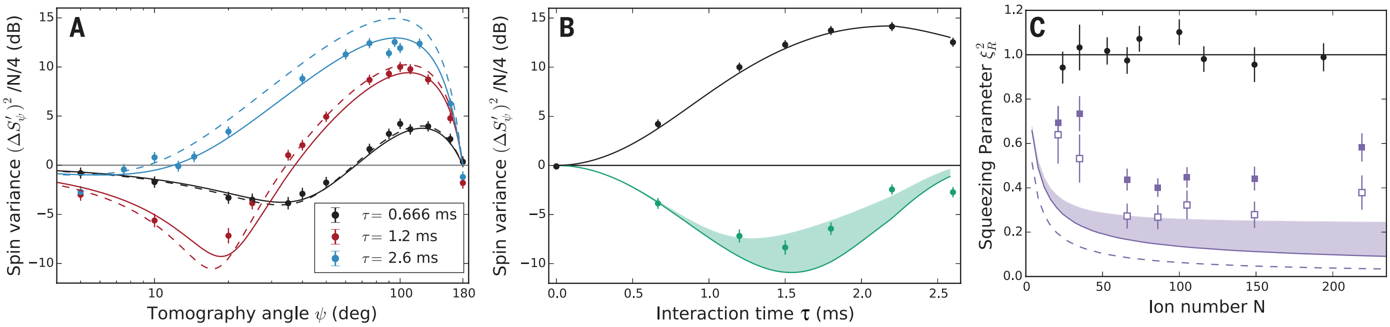


Fig. 3. Spin variance and entanglement. (**A**) Spin variance (symbols) as a function of tomography angle ψ for $N = 86 \pm 2$. The variance is calculated from 200 trials. The solid lines are a prediction, with no free parameters, assuming homogeneous Ising interactions and including decoherence from spontaneous emission (*28*). The dashed lines are theoretical predictions with the same interaction parameters but no decoherence. (**B**) The explicit time dependence of the spin variance for the ensemble in (A). The data for the minimum (green points) and maximum (black points) spin variance are shown with theory predictions of the optimal squeezing and antisqueezing (solid lines), including decoherence. Because our measurement of $(\Delta S'_\psi)^2$ has substantial granularity, we visualize the effect of finite sampling of ψ on the measured minimum variance using the green shaded region bounded by max $\{[\Delta S'_\psi(\psi_m \pm 5°)]^2\}$, where ψ_m corresponds to the angle that minimizes $(\Delta S'_\psi)^2$. The ±5° uncertainty does not have a visible effect in the squeezed component for short times, or for the antisqueezing component at any time. (**C**) Ramsey squeezing parameter measured for different ensemble sizes *N*. The black points show data for the initial unentangled spin state. The solid purple squares show the lowest directly measured ξ_R^2 with no corrections or subtractions of any detection noise for evaluation of the entanglement witness. The open squares show ξ_R^2 inferred by subtracting photon shot noise. The dashed line is the predicted optimal ξ_R^2 from coherent Ising interactions with no decoherence, and the solid line shows the limit including spontaneous emission assuming $\Gamma/\bar{J} = 0.05$, which is typical for our system. The shaded purple region accounts for finite sampling of ψ as in (B). All error bars denote SE.

time is 15 ms. Optical transitions to the $^2P_{3/2}$ states are used for state preparation, Doppler cooling, and projective measurement (*28*).

The Ising interaction is implemented by a spin-dependent optical dipole force (ODF) generated from the interference of a pair of detuned lasers (Fig. 1A). The ODF couples the spin and motional degrees of freedom through the interaction

$$\hat{H}_{\mathrm{ODF}} = \sum_{i=1}^{N} F_0 \cos(\mu t)\hat{z}_i\hat{\sigma}_i^z \tag{3}$$

where $\hat{z}_i$ is the position operator for ion *i*, $\mu/2\pi$ is the ODF laser beat frequency, and F_0 is the force amplitude, typically 30 yN. The ODF drives the axial drumhead modes of the planar ion crystal (*25*, *26*), generating an effective spin-spin interaction by modifying the ions' Coulomb potential energy (*29*). Detuning μ from ω_z changes the effective range of the spin-spin interaction $J_{i,j} \propto d_{i,j}^{-a}$, where d_{ij} is the ion separation. Although *a* can range from 0 to 3 (*24*), in this work we primarily drive the highest-frequency, center-of-mass (COM) mode at ω_z with ODF detunings $\delta = \mu - \omega_z$ ranging from about $2\pi \times 0.5$ kHz to $2\pi \times 3$ kHz, such that *a* varies from 0.02 to 0.18, respectively. The next closest axial motional mode frequency is more than $2\pi \times 20$ kHz lower than ω_z. Because $a << 1$, the Ising interaction is approximately independent of distance, resulting in a homogeneous pairwise coupling $J_{i,j} \approx \bar{J} = F_0^2/(4M\omega_z\delta)$, where *M* is the ion mass.

At the mean-field level, each spin precesses in an effective magnetic field determined by the couplings to other spins, described by the Hamiltonian

$$\hat{H}_{\mathrm{MF}} = \sum_{j=1}^{N} \frac{\bar{B}_j\hat{\sigma}_j^z}{2} \tag{4}$$

where

$$\bar{B}_j = \frac{2}{N}\sum_{i\neq j} J_{i,j}\langle\hat{\sigma}_i^z\rangle \tag{5}$$

We calibrated $\bar{J}$ through measurements of mean-field spin precession (*24*, *28*), typically finding $\bar{J}/\hbar \leq 3300\ \mathrm{s}^{-1}$. For the experiments described below, we started with all the spins initialized in an eigenstate of $\hat{\sigma}^x$ so that $\bar{B}_j = 0$. This choice of initial condition ensured that the observed physics are dominated by quantum correlations and decoherence alone.

State readout was performed using fluorescence from the Doppler cooling laser on the cycling transition (*28*). Ions in $|\uparrow\rangle$ fluoresce and ions in $|\downarrow\rangle$ are dark. Global fluorescence was collected with the side-view objective (Fig. 1A) and counted with a photomultiplier tube. We used the bottom-view image to count the number of ions and thereby calibrate the photon counts per ion (Fig. 1B). From the detected photon number, we could infer the bright-state population $N_\uparrow$, which is equivalent to a projective measurement of $\hat{S}_z = \hat{N}_\uparrow - (N/2)$, where $\hat{S}_z$ is the *z* component of the collective spin vector

$$\mathbf{S} = \frac{1}{2}\sum_{i}^{N} (\hat{\sigma}_i^x, \hat{\sigma}_i^y, \hat{\sigma}_i^z) \tag{6}$$

By performing a final global rotation before measuring, we could measure the moments of any component of **S**. The directly observed variance of the measurement $(\Delta S_z)^2$ is well described by the sum of two noise terms: spin noise $(\Delta S'_z)^2$ and photon shot noise $(\Delta S_{\mathrm{psn}})^2$. Here, ΔX indicates the standard deviation of repeated measurements of $\hat{X}$. In this paper, we use the underlying spin noise $(\Delta S'_z)^2 = (\Delta S_z)^2 - (\Delta S_{\mathrm{psn}})^2$ for comparison with theory predictions, but use the directly observed variance in the measurement $(\Delta S_z)^2$ for evaluating the spin-squeezing entanglement witness. The ratio $(\Delta S_{\mathrm{psn}})^2/(\Delta S'_z)^2$ is typically 0.13 (–8.8 dB), so the noise subtraction is small for all but the most squeezed states observed here. Other sources of state readout noise are not appreciable (*28*).

The depolarization of the collective spin length $|\langle\mathbf{S}\rangle|$, or contrast, caused by the Ising interaction is a canonical example of non-equilibrium quantum dynamics (*30*–*33*). Quantum correlations reduce the contrast and cause the collective spin state to wrap around the Bloch sphere that represents the state space (Fig. 2A). However, the contrast also decreases from decoherence, which destroys correlations, effectively shrinking the Bloch sphere. Our calculation accounts for both effects; for homogeneous Ising interactions $J_{i,j} = \bar{J}$ and at the time scales explored experimentally, the contrast is approximately (*28*) given by

$$|\mathbf{S}| \approx \exp(-\Gamma\tau)\frac{N}{2}\left[\cos\left(\frac{2\bar{J}}{N}\tau\right)\right]^{N-1} \tag{7}$$

where τ is the total ODF interaction time (Fig. 1C) and Γ is the total single-particle decoherence rate (*28*) due to spontaneous emission from the ODF lasers.

We show the depolarization dynamics of $|\langle\mathbf{S}\rangle|$ in our experiment in Fig. 2B, distinguishing effects of coherent interactions from decoherence. We determined $|\langle\mathbf{S}\rangle|$ from measurements of $\langle\hat{S}_x\rangle$, performing independent experiments to confirm that $\langle\hat{S}_y\rangle = \langle\hat{S}_z\rangle = 0$ after evolution under $\hat{H}_{\mathrm{I}}$. To distinguish the depolarization caused only by decoherence associated with the ODF lasers, we performed experiments at $\delta = +2\pi \times 50$ kHz, effectively eliminating the Ising coupling while leaving the spontaneous emission rate unchanged. The dashed line in Fig. 2B is a fit to the observed

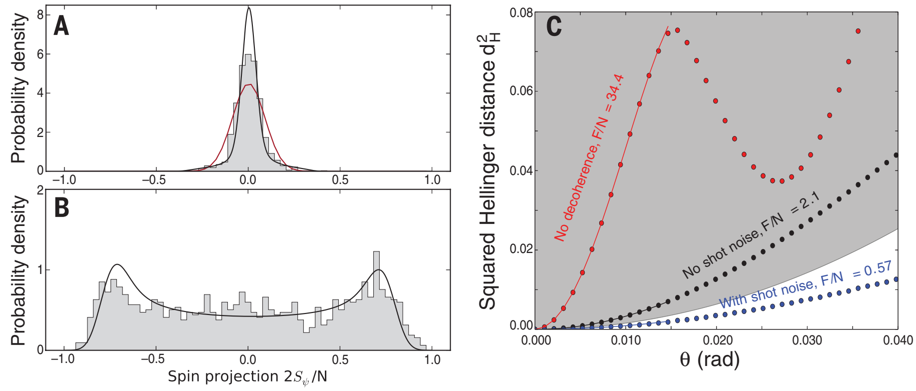


Fig. 4. Full counting statistics of a non-Gaussian spin state and theoretical Fisher information. (**A** and **B**) Histograms showing the squeezed (A) and antisqueezed (B) quadratures of the collective spin of $N = 127 \pm 4$ ions, corresponding to $\psi = 5.4°$ and $\psi = 92°$, respectively; $\tau = 3$ ms. The integral of the histogram is normalized to unity. The red line is the Gaussian distribution of the initial, unsqueezed state. The solid black line is the probability density predicted from numerical calculations (*28*), assuming homogeneous interactions and including decoherence from spontaneous emission and magnetic field fluctuations. We account for photon shot noise by convolving the theoretical probability density with a Gaussian distribution with a variance $(\Delta S_{\mathrm{psn}})^2/(N/4)$. (**C**) Extraction of the Fisher information from the theoretically computed Hellinger distance without (black) and with (blue) photon shot noise, including the effects of decoherence from spontaneous emission and magnetic field noise. Shown in red is the Hellinger distance in the absence of decoherence or photon shot noise, for comparison. The points denote computed values of the Hellinger distance; the lines are small-angle quartic fits. The gray swath denotes the region of entangled states (*28*).

exponential decay, measuring Γ in our system (*28*). The faster contrast decay for μ tuned near ω_z is in good agreement with Eq. 7 for a range of system sizes. For these data, $\delta = 4\pi/\tau$, ensuring spin-motion decoupling of the COM mode at the end of the experiment (*25*). The collapse of the data to a single curve when plotted as a function of $2\bar{J}\tau/\sqrt{N}$ (Fig. 2B, inset) provides strong evidence that the depolarization is primarily the result of spin-spin interactions. However, depolarization dynamics alone are not enough to prove that entanglement exists in the ensemble.

To verify entanglement, we used the Ramsey squeezing parameter ξ_R^2, which only requires measuring the variance of collective observables, instead of full state tomography. The Ramsey squeezing parameter is

$$\xi_R^2 = N\frac{\min_\psi[(\Delta S_\psi)^2]}{|\langle \mathbf{S}\rangle|^2} \tag{8}$$

where

$$\hat{S}_\psi = \frac{1}{2}\sum_i^N \cos(\psi)\hat{\sigma}_i^z + \sin(\psi)\hat{\sigma}_i^y \tag{9}$$

and $\min_\psi$ indicates taking the minimum as a function of ψ. For an unentangled spin state, polarized along the x axis, $|\langle \mathbf{S}\rangle| = N/2$ and the spin noise is set by Heisenberg uncertainty relations to $(\Delta S_y)^2 = (\Delta S_z)^2 = N/4$, so $\xi_R^2 = 1$. This quantum noise limits the signal-to-noise ratio for a wide range of quantum sensors based on ensembles of independent quantum objects (*34*). Nonclassical correlations can redistribute quantum noise between two orthogonal quadratures of the collective spin, squeezing the noise such that $(\Delta S_\psi)^2 < N/4$ and $\xi_R^2 < 1$. These squeezed states are entangled (*35*), and furthermore, $\xi_R^2 < 1$ proves that the entangled state is a useful resource for precision sensing (*15*, *34*–*43*).

At short times, the non-equilibrium spin dynamics caused by the Ising interaction can produce spin-squeezed states (*15*, *32*, *34*, *44*). Figure 3, A and B, shows the measured time evolution of the spin variance $(\Delta S_\psi)^2$ of 86 ions, normalized to the spin variance of the initial, unentangled state. We compared the data to an analytic model (*32*) that assumes homogeneous Ising interactions and fully accounts for both elastic and spin-changing spontaneous emission. The data clearly show the development of squeezed and antisqueezed quadratures, and deviations from perfectly coherent Ising dynamics are well described by the effects of spontaneous emission alone. Similar data for different values of N are shown in (*28*).

Using measurements of the directly observed spin variance $(\Delta S_\psi)^2$ and contrast $|\langle \mathbf{S}\rangle|$, we obtained ξ_R^2 for a range of values of τ. As shown by a plot of the minimum observed ξ_R^2 for each N in Fig. 3C, the entanglement witness $\xi_R^2 < 1$ is satisfied for seven independent data sets with N ranging from 21 to 219. We observe a minimum $\xi_R^2 = -4.0 \pm 0.9$ dB for $N = 84$ ions. We also show ξ_R^2 measured for the initial state, confirming our calibration of N. For comparison, Fig. 3C shows the absolute minimum ξ_R^2 predicted for coherent Ising interactions. The majority of the observed discrepancy for ensembles ranging from 60 to 150 ions is accounted for by photon shot noise, spontaneous emission, and the finite sampling of τ and ψ. For other ion numbers (*28*), we still observe good agreement in the antisqueezed spin variance, but the minimum spin variance and ξ_R^2 deviate further from the prediction. We attribute the deviation to technical noise sources (*28*).

The Ramsey squeezing parameter is an effective entanglement witness at short times when quantum noise is approximately Gaussian. At longer times, the growth of spin correlations causes both the depolarization seen in Fig. 2 and the increase in $\min_\psi[(\Delta S_\psi)^2]$, due to the appearance of non-Gaussian quantum noise in the collective spin. Both effects cause ξ_R^2 to increase above 1, which we call an oversqueezed state. Oversqueezed states can be entangled (*45*); however, ξ_R^2 can also increase simply because of decoherence.

Signatures of quantum correlations at longer interaction times are seen in a histogram of the measurements of $\langle \hat{S}_\psi\rangle$ for an oversqueezed state of 127 ions after an interaction time of $\tau = 3$ ms (Fig. 4B). For times well beyond the optimum squeezing time, we see a clear non-Gaussian distribution for the antisqueezed quadrature. The distribution for $\psi = 5.4°$ (Fig. 4A) also contains non-Gaussian characteristics in the tails away from the narrow central feature. We found good agreement between these data and a theoretical model of the full counting statistics. Even though $\xi_R^2 = 26$, the theoretically predicted state is entangled, as shown by an entanglement witness based on the Fisher information F (Fig. 4C). The quantum Fisher information has been used as an entanglement witness in other trapped-ion simulators (*46*). Here, we bound the value of F using the approach in (*40*) and find $F/N > 2.1$, which satisfies the inequality of the entanglement witness $F/N > 1$ (*45*). Photon shot noise in our measurement limits our capability to directly witness the entanglement experimentally (*28*), but the good agreement with theory indicates that the state of the ensemble is consistent with an entangled, oversqueezed state. Additionally, we experimentally confirmed that this procedure for bounding F witnesses entanglement of squeezed states (fig. S11). The full counting statistics are only efficiently computable for homogeneous couplings, a good approximation for the small detunings δ considered here. For future work with inhomogeneous Ising coupling, obtaining the full counting statistics theoretically will likely be intractable for more than 20 to 30 spins.

The techniques presented here are applicable to precision sensors using trapped ions, where the number of ions is limited by systematic errors arising from ion motion (*47*), and could be useful for quantum-enhanced metrology with non-Gaussian spin states (*40*, *48*–*50*). These results benchmark controlled quantum evolution in a 2D platform with more than 200 spins, establishing a foundation for future experiments studying the full transverse-field Ising model in regimes inaccessible to classical computation. With the implementation of single-spin readout, the simulator could provide unique opportunities to study the dynamics of spin correlations in 2D systems,

such as Lieb-Robinson bounds (*19*) and many-body localization in the presence of disorder (*5*, *6*).

REFERENCES AND NOTES

1. R. P. Feynman, *Int. J. Theor. Phys.* **21**, 467–488 (1982).
2. P. Hauke, F. M. Cucchietti, L. Tagliacozzo, I. Deutsch, M. Lewenstein, *Rep. Prog. Phys.* **75**, 082401 (2012).
3. I. M. Georgescu, S. Ashhab, F. Nori, *Rev. Mod. Phys.* **86**, 153–185 (2014).
4. R. J. Elliott, P. Pfeuty, C. Wood, *Phys. Rev. Lett.* **25**, 443–446 (1970).
5. K. Binder, A. P. Young, *Rev. Mod. Phys.* **58**, 801–976 (1986).
6. R. Nandkishore, D. A. Huse, *Annu. Rev. Condens. Matter Phys.* **6**, 15–38 (2015).
7. D. Belitz, T. R. Kirkpatrick, T. Vojta, *Rev. Mod. Phys.* **77**, 579–632 (2005).
8. S. Sachdev, *Quantum Phase Transitions* (Wiley, 2007).
9. T. F. Rønnow *et al.*, *Science* **345**, 420–424 (2014).
10. A. Lucas, *Front. Phys.* **2**, 5 (2014).
11. J. Simon *et al.*, *Nature* **472**, 307–312 (2011).
12. A. de Paz *et al.*, *Phys. Rev. Lett.* **111**, 185305 (2013).
13. T. Fukuhara *et al.*, *Phys. Rev. Lett.* **115**, 035302 (2015).
14. R. C. Brown *et al.*, *Science* **348**, 540–544 (2015).
15. I. D. Leroux, M. H. Schleier-Smith, V. Vuletić, *Phys. Rev. Lett.* **104**, 073602 (2010).
16. B. Yan *et al.*, *Nature* **501**, 521–525 (2013).
17. D. Porras, J. I. Cirac, *Phys. Rev. Lett.* **92**, 207901 (2004).
18. P. Jurcevic *et al.*, *Nature* **511**, 202–205 (2014).
19. P. Richerme *et al.*, *Nature* **511**, 198–201 (2014).
20. R. Islam *et al.*, *Science* **340**, 583–587 (2013).
21. R. Schmied, J. H. Wesenberg, D. Leibfried, *Phys. Rev. Lett.* **102**, 233002 (2009).
22. C. D. Bruzewicz, R. McConnell, J. Chiaverini, J. M. Sage, http://arxiv.org/abs/1511.03293 (2015).
23. J. D. Baltrusch, A. Negretti, J. M. Taylor, T. Calarco, *Phys. Rev. A* **83**, 042319 (2011).
24. J. W. Britton *et al.*, *Nature* **484**, 489–492 (2012).
25. B. C. Sawyer *et al.*, *Phys. Rev. Lett.* **108**, 213003 (2012).
26. C.-C. J. Wang, A. C. Keith, J. K. Freericks, *Phys. Rev. A* **87**, 013422 (2013).
27. M. J. Biercuk *et al.*, *Quantum Inf. Comput.* **9**, 920–949 (2009).
28. See supplementary materials on *Science* Online.
29. D. Leibfried *et al.*, *Nature* **422**, 412–415 (2003).
30. M. Kastner, *Phys. Rev. Lett.* **106**, 130601 (2011).
31. M. V. D. Worm, B. C. Sawyer, J. J. Bollinger, M. Kastner, *New J. Phys.* **15**, 083007 (2013).
32. M. Foss-Feig, K. R. A. Hazzard, J. J. Bollinger, A. M. Rey, *Phys. Rev. A* **87**, 042101 (2013).
33. K. R. A. Hazzard *et al.*, *Phys. Rev. A* **90**, 063622 (2014).
34. D. Wineland, J. Bollinger, W. Itano, F. Moore, D. J. Heinzen, *Phys. Rev. A* **46**, R6797 (1992).
35. A. Sørensen, L. M. Duan, J. I. Cirac, P. Zoller, *Nature* **409**, 63–66 (2001).
36. L. K. Shalm, R. B. A. Adamson, A. M. Steinberg, *Nature* **457**, 67–70 (2009).
37. W. Wasilewski *et al.*, *Phys. Rev. Lett.* **104**, 133601 (2010).
38. C. D. Hamley, C. S. Gerving, T. M. Hoang, E. M. Bookjans, M. S. Chapman, *Nat. Phys.* **8**, 305–308 (2012).
39. C. F. Ockeloen, R. Schmied, M. F. Riedel, P. Treutlein, *Phys. Rev. Lett.* **111**, 143001 (2013).
40. H. Strobel *et al.*, *Science* **345**, 424–427 (2014).
41. N. Behbood *et al.*, *Phys. Rev. Lett.* **113**, 093601 (2014).
42. O. Hosten, N. J. Engelsen, R. Krishnakumar, M. A. Kasevich, *Nature* **529**, 505–508 (2016).
43. K. C. Cox, G. P. Greve, J. M. Weiner, J. K. Thompson, *Phys. Rev. Lett.* **116**, 093602 (2016).
44. M. Kitagawa, M. Ueda, *Phys. Rev. A* **47**, 5138–5143 (1993).
45. L. Pezzé, A. Smerzi, *Phys. Rev. Lett.* **102**, 100401 (2009).
46. J. Smith *et al.*, http://arxiv.org/abs/1508.07026 (2015).
47. K. Arnold *et al.*, *Phys. Rev. A* **92**, 032108 (2015).
48. B. Lücke *et al.*, *Phys. Rev. Lett.* **112**, 155304 (2014).
49. F. Haas, J. Volz, R. Gehr, J. Reichel, J. Estève, *Science* **344**, 180–183 (2014).
50. R. McConnell, H. Zhang, J. Hu, S. Ćuk, V. Vuletić, *Nature* **519**, 439–442 (2015).

ACKNOWLEDGMENTS

We thank D. Hume, J. Thompson, A. Keith, J. Schachenmayer, A. Safavi-Naini, M. Gaerttner, and M. Kastner for helpful discussions. Supported by NSF grant PHY 1521080, JILA-NSF grant PFC-1125844, the Army Research Office, and the Air Force Office of Scientific Research and its Multidisciplinary University Research Initiative (A.M.R.) and by a National Research Council Research Associateship Award at NIST (J.G.B., M.F.-F., and M.L.W.). All authors acknowledge financial support from NIST. This manuscript is the contribution of NIST and is not subject to U.S. copyright.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1297/suppl/DC1
Materials and Methods
Figs. S1 to S11
References (*51*–*65*)

7 December 2015; accepted 4 May 2016
10.1126/science.aad9958

OPTICAL MATERIALS

A highly efficient directional molecular white-light emitter driven by a continuous-wave laser diode

Nils W. Rosemann,[1,2] Jens P. Eußner,[2,3] Andreas Beyer,[1,2] Stephan W. Koch,[1,2] Kerstin Volz,[1,2] Stefanie Dehnen,[2,3]* Sangam Chatterjee[1,2,4]*

Tailored light sources have greatly advanced technological and scientific progress by optimizing the emission spectrum or color and the emission characteristics. We demonstrate an efficient spectrally broadband and highly directional warm-white-light emitter based on a nonlinear process driven by a cheap, low-power continuous-wave infrared laser diode. The nonlinear medium is a specially designed amorphous material composed of symmetry-free, diamondoid-like cluster molecules that are readily obtained from ubiquitous resources. The visible part of the spectrum resembles the color of a tungsten-halogen lamp at 2900 kelvin while retaining the superior beam divergence of the driving laser. This approach of functionalizing energy-efficient state-of-the-art semiconductor lasers enables a technology complementary to light-emitting diodes for replacing incandescent white-light emitters in high-brilliance applications.

The impact of well-managed light on our everyday life is immeasurable (*1*, *2*). The light-emitting diode (LED) is one of the most prominent developments since the invention of incandescent lightbulbs in the late 1800s (*3*). The latter dissipate most energy in the infrared as heat, whereas typical white LEDs cover only the visible spectrum. Most prominent examples of white-light LEDs are based on gallium nitride (*4*, *5*). Their narrow-band ultraviolet (UV) emission is converted into visible light by applying phosphors (*6*–*9*). This cold light has tremendous advantages with respect to energy efficiency. Other concepts pursued for efficient white-light generation include the combination of red, green, and blue emitters (*10*), which is currently the path of choice for organic LEDs (*11*–*14*). All types of LEDs excel due to their virtually Lambertian emission patterns that are highly desirable for applications like active displays that require large viewing angles (*15*). However, this poses challenges in targeted illumination and projection of light due to the associated large etendue $G = A\Omega$, where A is the source area and Ω is the solid angle of emission (*16*). Ideally, the etendue remains constant throughout an entire optical system where light undergoes perfect reflections or refractions. It can increase—for example, when impinging on a diffusor—but cannot be decreased without loss in radiance. This renders low-etendue sources extremely desirable for devices requiring high spatial resolution like microscopes or for applications with high throughput, such as projection systems.

Other concepts of white-light generation by monochromatic sources besides phosphors rely on nonlinear effects that provide very broadband supercontinua and are widely used in many scientific applications (*17*, *18*). These are often referred to as brilliant sources. They generally feature small, point-source–like emission areas due to the tightly focused short-pulsed driving lasers that are used to overcome the vast peak electric field strength required to invoke the extremely nonlinear effects such as soliton formation (*19*). Hence, the related challenges, such as the system size, price, and energy requirements, restrict the use of supercontinuum sources to scientific laboratory use and the medical sector—for example, in coherent anti–Stokes Raman scattering (*20*) or optical coherence tomography (*21*), as well as for defense and security applications.

Here, we use a molecule-based solid compound as an extremely nonlinear medium. It enables the steady-state operation of a low-etendue, directional broadband white-light source covering the entire

[1]Fachbereich Physik, Philipps-Universität Marburg, DE-35032 Marburg, Germany. [2]Wissenschaftliches Zentrum für Materialwissenschaften, Philipps-Universität Marburg, DE-35032 Marburg, Germany. [3]Fachbereich Chemie, Philipps-Universität Marburg, DE-35032 Marburg, Germany. [4]Institute of Experimental Physics I, Justus-Liebig-University, DE-35392 Giessen, Germany.
***Corresponding author. Email: stefanie.dehnen@uni-marburg.de (S.D.); sangam.chatterjee@physik.uni-giessen.de (S.C.)**

visible spectrum driven by a low-cost infrared laser diode. The compound comprises semiconductor-based cluster molecules decorated with covalently attached organic ligands supplying quasidelocalized electrons. The overall goal was to synthesize an amorphous compound combining a suitable bandgap inorganic semiconductor cluster core with organic ligands providing delocalized electrons in the electronic ground state of the molecule, while being composed of components that are ubiquitous, thus readily obtainable and cheap. We produced tin-sulfide–based molecules with an adamantane-like, thus diamondoid, $[Sn_4S_6]$ scaffold. The latter is free of inversion symmetry because it has a tetrahedral shape. Lower molecular symmetry and delocalization of electronic states are realized by decoration of the core with randomly oriented organic ligands $R^{deloc} = 4–(CH_2{=}CH)–C_6H_4$ (Fig. 1A). The steric influence of the organic ligands defines the molecular structure of the cluster core (*22*–*24*) and the noncrystalline nature of the compound. It prevents polymerization of the inorganic cluster moieties into the binary SnS_2 solid; the vinyl groups in para position of the tin atoms are also available for further chemical modification (*25*). Additionally, they enable covalent attachment of the clusters to inorganic materials.

The compound $[(R^{deloc}Sn)_4S_6]$ was obtained as a fine amorphous powder (Fig. 1B). Its identity was confirmed by mass spectrometry (*26*) (fig. S1). The compound is nonvolatile, air-stable, and thermally stable up to 573 K (*26*) (fig. S2). Its molecular structure was rationalized by quantum chemical calculations employing density-functional-theory (DFT) methods (*26*) (fig. S3). The compound retains its chemical and physical characteristics when embedded inside an acrylamide-based matrix. The emission for 800-nm continuous-wave (CW) laser excitation gives a warm-white color impression (Fig. 1C). It is very close to a standard tungsten-halogen light source at 2856 K for the maximum pump fluency and changes with variation of the excitation density (Fig. 1D).

Although the color impression closely resembles an incandescent source, the characteristic directional features of the driving CW laser are retained by the nonlinear medium. The angular emission pattern for excitation with a loosely focused CW near-infrared laser beam (Fig. 2A) (*26*) shows that the white-light emission features a very narrow angular spread. Its emission cone is broader than that of the driving laser, as expected for scattering due to the amorphous character of the compound (*26*) (fig. S6). This implies that our approach should enable the realization of high-brilliance sources for targeted, directional illumination and projection applications. Here, one should hence be able to outperform conventional approaches with respect to directionality: Thermal emitters are inherently omnidirectional, and LEDs typically exhibit Lambertian emission (*16*).

The emission spectra corresponding to the color temperatures in Fig. 1D are given in Fig. 2B. The dispersed emission covers the entire visible spectrum; its spectral weight is shifted toward lower energies compared with characteristic white-light LED emission. The spectral distribution of the white light is virtually independent of the excitation wavelength in the range from 725 to 1050 nm (*26*) (fig. S4). This property is highly desirable for integration into laser-diode devices because it implies robustness to thermal- or manufacturing-related variations in the driving-laser wavelength. The nonlinear behavior is visualized in the input-output characteristic given in Fig. 2C. The white-light output power as a function of pump-power density reveals an extreme nonlinearity, which scales approximately to the eighth power. Currently, the optimum supercontinuum generation efficiency close to the destruction threshold is determined to be in the range of 10% (*26*). Even at this early stage and imperfect sample quality, this efficiency is comparable to widely used phosphors. The samples have shown notable long-term stability under operation conditions for several months.

To gain insight into the underlying mechanism of white-light generation, we compare the spontaneous emission for above-bandgap UV excitation to the white-light characteristics (Fig. 3A). The spontaneous emission is a mirror image of the linear absorption, as expected from the Franck-Condon principle (*27*). The white-light spectrum is shifted to lower energies; only the high-energy cutoff appears to be limited by reabsorption in the dense amorphous molecular solid. Furthermore, it should be noted that the spontaneous emission is several orders of magnitude less bright than the white light. The photon energy of the driving laser is detuned very far off resonance, and no indications of emission after multiphoton excitation are seen. Further differences are observed for the lifetimes for pulsed excitation of the spontaneous emission and the white light (Fig. 3B). The spontaneous emission decays on a 100-ps time scale, whereas the white light

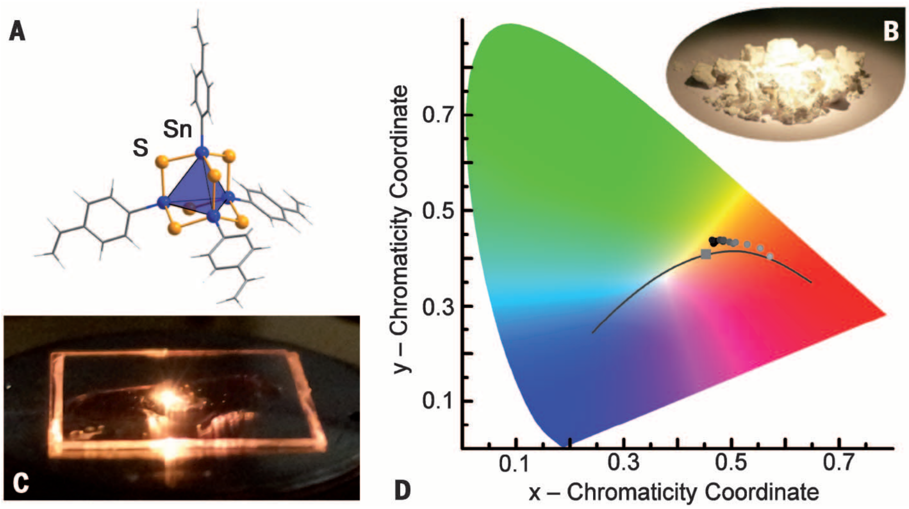


Fig. 1. Molecular structure and appearance as well as color temperatures associated with the emission. (**A**) Adamantane-like cluster $[(R^{deloc}Sn)_4S_6]$ ($R^{deloc} = 4–(CH_2{=}CH)–C_6H_4$), with tin and sulfur atoms drawn as blue and yellow spheres, respectively; carbon (gray) and hydrogen (white) atoms are given as wires. (**B**) Photograph of the as-prepared powder. (**C**) Photograph of a polymer film containing the cluster sandwiched between two cover glass slips excited by 800-nm laser light in the bright center spot. (**D**) Color temperatures given for various excitation fluencies, as indicated by individual gray-scale data points. The characteristic ideal black-body emission for various temperatures is indicated by the solid line; the square indicates the color temperature of standard emitter at T = 2856 K.

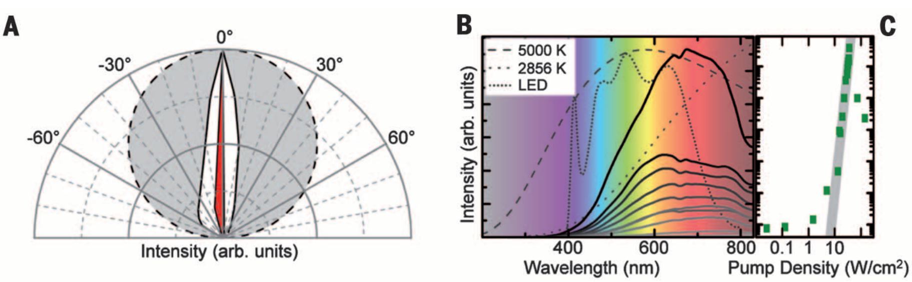


Fig. 2. Emission characteristics. (**A**) Highly directional spatial emission pattern of the white-light spectrum (white) and the CW excitation laser at 980 nm (red). The intensity distribution of a perfect Lambertian emitter (gray) is given for reference. (**B**) White-light spectra for a pump wavelength of 980 nm. The pump power is varied from 6 mW (light gray solid line) to 18 mW (black solid line). The normalized curves for black-body radiation (T = 5000 K, dashed line; T = 2856 K, spaced dots) and a GaN-based white-light LED (narrow dots) are shown for comparison. (**C**) Double-logarithmic plot of the white light input-output characteristics.

intriguingly shows virtually no time dynamics at all, inferring an instrument-limited lifetime of more than 10 μs. The long lifetime excludes conventional coherent processes as the source of the broad spectrum, and it implies that for CW irradiation, the directionality is caused by a phased-array effect induced by the continuously present electric field of the driving laser.

All of the above considerations imply a mechanism for white-light generation (Fig. 3C), whereby the near-infrared laser drives the virtually delocalized electrons supplied by the π-electron systems of the organic ligands. For CW irradiation, the charges are driven in the electronic ground-state potential of the molecule and dominantly relax via radiative loss in energy, commonly termed bremsstrahlung. These features are captured qualitatively in a model of the emission resembling the classical motion of an electron invoked by an external driving field (*26*). The ground-state potential landscape is approximated by an anharmonic oscillator with a third-order perturbation. This simplification is justified by the small energy scale below 3.5 eV considered here, compared with the ground-state ionization threshold calculated to 13.46 eV (*26*). Results from the calculation are plotted in Fig. 3D. The calculation yields the dominant emission peak around 2 eV in the visible, excellently reproducing the experimental data. The near-infrared emission peak predicted by the calculations is also observed experimentally, corroborating the applicability of this phenomenological model.

The reemission during the accelerated motion in the anharmonic molecular potential is conceptually similar to high-harmonic generation in noble gases or optically driven gas plasmas (*28*, *29*) and the resulting plateau formation (*30*). The most remarkable difference, however, is the involvement of electrons in the electronic ground state only. This infers the low electric field strength required to invoke the broadband emission accessible even for CW lasers. Hence, the nonlinearity should depend critically on the electrons available in conjugated π-systems of the ligands and on the composition of the cluster core that provides the high-energy cutoff (Fig. 3A). This explanation is supported by systematic investigations with exchanged core and ligand structures. Replacing Sn with Ge increases both the ground-state potential depth and the fundamental electronic transition energies. Consequently, the supercontinuum's spectral bandwidth is increased toward higher energies. The importance of amorphousness becomes evident when replacing Sn with Si; a crystalline material is obtained and the supercontinuum is quenched.

The clusters are potentially integrated into a monolithic device. During vacuum deposition, the clusters form thin amorphous layers on hydrogen-terminated silicon single crystals and on GaAs, the latter being particularly important for integration into diode lasers. The high-angle annular dark-field scanning transmission electron microscopy (TEM) image for deposition on a GaAs (001) surface is given in Fig. 4A (*26*) [fig. S8 for the deposition on Si (001)]. The width of the amorphous molecular layer observed in the high-resolution micrograph (Fig. 4B) is in excellent agreement with a cluster molecule's dimensions; this suggests that the surface is coated by a self-limited monolayer. The composition across the interface (Fig. 4C) and the energy-dispersive x-ray (EDX) (Fig. 4A) corroborate these findings.

Our approach provides a route for a directional white-light device for low-etendue applications complementary to conventional nonlinear

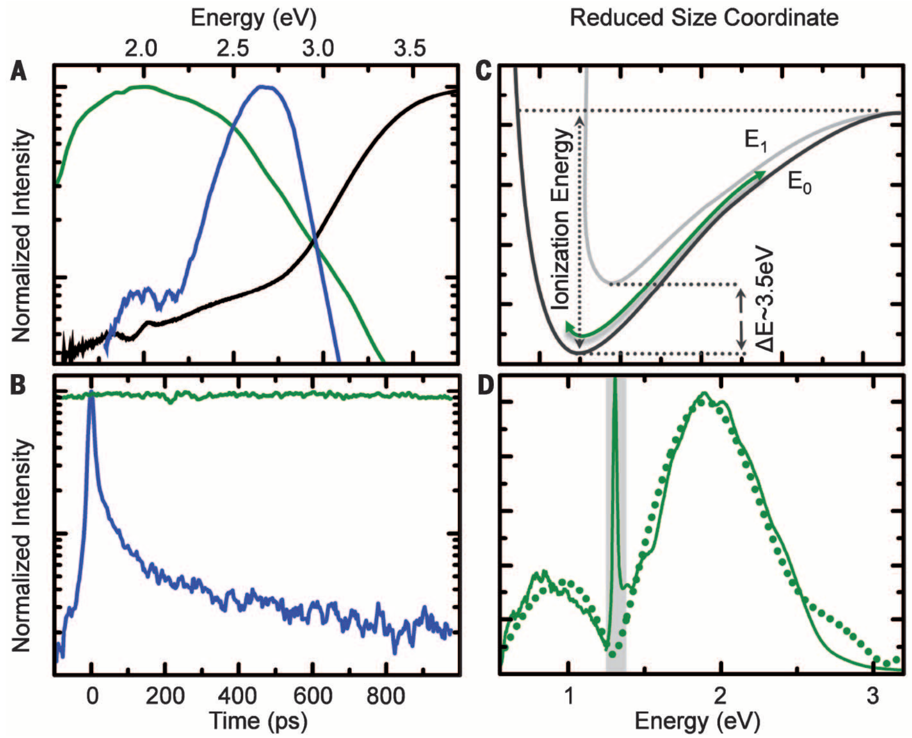


Fig. 3. Spectral and temporal emission characteristics and mechanism. (**A**) Normalized linear absorption spectrum (black), spontaneous emission for UV excitation above the fundamental electronic transition energy (blue), and white-light emission spectrum (green) for the driving infrared laser on a semilogarithmic scale. (**B**) The spontaneous emission (blue) decays significantly faster than the white light (green). (**C**) Schematic illustration portraying the white-light emission due to the accelerated motion of an electron (indicated by green trajectory) in the anharmonic electronic ground-state potential (solid curve). (**D**) Experimental (solid) and calculated (dotted) white-light emission spectra agree excellently; the scattered part of the driving laser (gray-shaded area) is not included in the simulation.

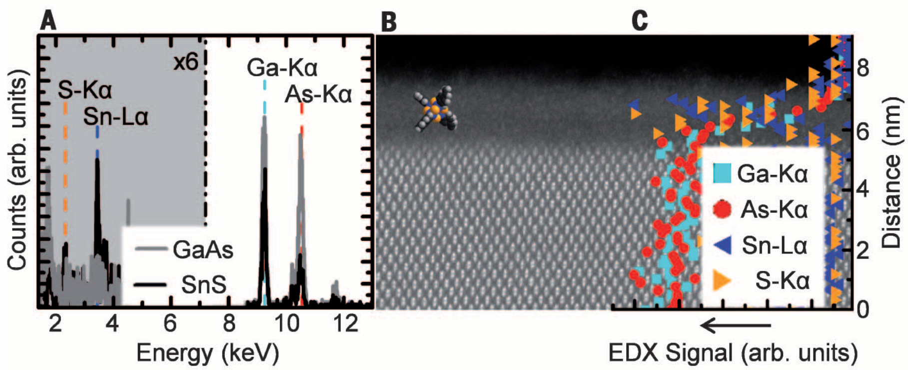


Fig. 4. Compositional and structural characterization of functionalization characteristics on GaAs. (**A**) EDX spectra revealing the contributions of Sn and S in the amorphous cluster layer and Ga and As in the crystalline substrate. (**B**) The self-assembled monolayer shows long-range homogeneity and lacks any observable structure, and hence is perfectly amorphous, as can be seen from the high-resolution micrograph; a scaled structure model is overlaid on the micrograph to illustrate the size. (**C**) Overlaid at the right side of the micrograph: EDX line scans indicating the distribution of the constituents. The length scale for both (B) and (C) is defined by the right vertical axis.

sources or solid-state emitters such as LEDs by the functionalization of a low-cost infrared diode laser.

REFERENCES AND NOTES

1. Committee on Optical Science and Engineering, Commission on Physical Sciences, Mathematics, and Applications, Division on Engineering and Physical Sciences, National Research Council, *Harnessing Light: Optical Science and Engineering for the 21st Century* (National Academy Press, Washington, DC, 1998).
2. E. F. Schubert, J. K. Kim, *Science* **308**, 1274–1278 (2005).
3. S. Pimputkar, J. S. Speck, S. P. DenBaars, S. Nakamura, *Nat. Photonics* **3**, 180–182 (2009).
4. S. Nakamura, T. Mukai, M. Senoh, *Appl. Phys. Lett.* **64**, 1687–1689 (1994).
5. S. Nakamura, *Rev. Mod. Phys.* **87**, 1139–1151 (2015).
6. C. C. Yang *et al.*, *Appl. Phys. Lett.* **90**, 123503 (2007).
7. C. C. Lin, R.-S. Liu, *J. Phys. Chem. Lett.* **2**, 1268–1277 (2011).
8. N. C. George, K. A. Denault, R. Seshadri, *Annu. Rev. Mater. Res.* **43**, 481–501 (2013).
9. M. Shang, C. Li, J. Lin, *Chem. Soc. Rev.* **43**, 1372–1386 (2014).
10. S. Muthu, F. J. Schuurmans, M. D. Pashley, *IEEE J. Sel. Top. Quantum Electron.* **8**, 333–338 (2002).
11. J. H. Burroughes *et al.*, *Nature* **347**, 539–541 (1990).
12. M. A. Baldo *et al.*, *Nature* **395**, 151–154 (1998).
13. T. Tsujimura, *OLED Display Fundamentals and Applications* (Wiley Publishing, Hoboken, NJ, 2012).
14. N. Thejokalyani, S. J. Dhoble, *Renew. Sustain. Energy Rev.* **32**, 448–467 (2014).
15. J. S. Kim, P. K. Ho, N. C. Greenham, R. H. Friend, *J. Appl. Phys.* **88**, 1073–1081 (2000).
16. J. M. Palmer, B. G. Grant, *The Art of Radiometry* (SPIE Press, Bellingham, WA, 2009).
17. R. R. Alfano, *The Supercontinuum Laser Source* (Springer-Verlag, New York, 2016).
18. J. M. Dudley, G. Genty, S. Coen, *Rev. Mod. Phys.* **78**, 1135–1184 (2006).
19. S. Coen *et al.*, *Opt. Lett.* **26**, 1356–1358 (2001).
20. H. Kano, H. O. Hamaguchi, *Appl. Phys. Lett.* **85**, 4298–4300 (2004).
21. B. Povazay *et al.*, *Opt. Lett.* **27**, 1800–1802 (2002).
22. H. Berwe, A. Haas, *Chem. Ber.* **120**, 1175–1182 (1987).
23. Z. Hassanzadeh Fard, M. R. Halvagar, S. Dehnen, *J. Am. Chem. Soc.* **132**, 2848–2849 (2010).
24. J. P. Eußner, S. Dehnen, *Chem. Commun.* **50**, 11385–11388 (2014).
25. W. Zhang, J. L. Loebach, S. R. Wilson, E. N. Jacobsen, *J. Am. Chem. Soc.* **112**, 2801–2803 (1990).
26. Materials and methods are available as supplementary materials on *Science* Online.
27. P. W. Atkins, R. S. Friedman, *Molecular Quantum Mechanics* (Oxford Univ. Press, Oxford, ed. 4, 2004).
28. K. J. Schafer, B. Yang, L. F. DiMauro, K. C. Kulander, *Phys. Rev. Lett.* **70**, 1599–1602 (1993).
29. P. B. Corkum, *Phys. Rev. Lett.* **71**, 1994–1997 (1993).
30. G. G. Paulus, W. Nicklich, H. Xu, P. Lambropoulos, H. Walther, *Phys. Rev. Lett.* **72**, 2851–2854 (1994).

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft within the framework of the research training group GRK1782. N.W.R., S.W.K., and S.C. thank U. Huttner and M. Kira for stimulating discussions on the model.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1301/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S3
References (*31*–*46*)

3 March 2016; accepted 12 May 2016
10.1126/science.aaf6138

ORGANIC CHEMISTRY

Native functionality in triple catalytic cross-coupling: sp^3 C–H bonds as latent nucleophiles

Megan H. Shaw,* Valerie W. Shurtleff,* Jack A. Terrett,* James D. Cuthbertson, David W. C. MacMillan†

The use of sp^3 C–H bonds—which are ubiquitous in organic molecules—as latent nucleophile equivalents for transition metal–catalyzed cross-coupling reactions has the potential to substantially streamline synthetic efforts in organic chemistry while bypassing substrate activation steps. Through the combination of photoredox-mediated hydrogen atom transfer (HAT) and nickel catalysis, we have developed a highly selective and general C–H arylation protocol that activates a wide array of C–H bonds as native functional handles for cross-coupling. This mild approach takes advantage of a tunable HAT catalyst that exhibits predictable reactivity patterns based on enthalpic and bond polarity considerations to selectively functionalize α-amino and α-oxy sp^3 C–H bonds in both cyclic and acyclic systems.

Over the past 50 years, transition metal–catalyzed cross-coupling reactions have transformed the field of synthetic organic chemistry via the evolution of a wide variety of C–C and C–heteroatom bond-forming reactions (*1*, *2*). During this time, the seminal studies of Negishi, Suzuki, Miyaura, Stille, Kumada, and Hiyama have inspired numerous protocols to construct carbon–carbon bonds using palladium, nickel, or iron catalysis. These strategies enable highly efficient and regiospecific fragment couplings with high functional group tolerance, facilitating the application of modular building blocks in early- or late-stage synthetic efforts. Traditionally, cross-coupling methods have relied on the use of organometallic nucleophiles such as aryl or vinyl boronic acids, zinc halides, stannanes, or Grignard reagents that undergo addition to a corresponding metal-activated aryl or vinyl halide.

An emerging strategy for C–C bond formation has been the application of native organic functionality as latent nucleophilic handles for transition metal–mediated cross-couplings. In this context, the use of olefin, methoxy, acetoxy, and carboxylic acid moieties as organometallic replacements has enabled a variety of carbon–carbon bond formation protocols using feedstock materials (*3*–*8*). However, the most common approach for transition metal–mediated native functionalization has been the use of C–H bonds—the most ubiquitous chemical bonds found in nature—as nucleophilic coupling partners. Among the well-established challenges with sp^3 C–H bond functionalization, regioselectivity is perhaps preeminent, given that organic molecules incorporate a diverse combination of methyl, methylene, and/or methine groups. Several elegant methodologies have navigated this question via the use of directing groups to accomplish selective sp^3 C–H bond functionalization (*9*–*13*), or more recently by focusing on the use of inductive effects to deactivate C–H bonds (*14*). Enzymes accomplish selective sp^3 C–H bond functionalization by taking advantage of the diverse electronic and enthalpic characteristics of carbon–hydrogen bonds found within complex organic molecules (*15*). Inspired by this biochemical blueprint, we speculated that a small-molecule catalyst platform could be developed that would differentiate between a diverse range of C–H groups using a combination of bond energies and polarization, thereby enabling a unique pathway toward native arylation or vinylation.

A fundamental mechanistic step in organic synthesis is the simultaneous movement of a proton and an electron—a process termed hydrogen atom transfer (HAT) (*16*, *17*). HAT has long served as an effective way to access radical intermediates in organic chemistry; however, the capacity to regioselectively abstract hydrogens among a multitude of diverse C–H locations has been notoriously difficult to control. Recently, driven by developments in small-molecule catalyst design, general methods for C–H bond functionalization via HAT have begun to achieve levels of selectivity that were previously restricted to enzymatic systems (*18*, *19*). In this context, our laboratory has demonstrated that photoredox-mediated HAT catalysis can exploit native sp^3 C–H bonds for a range of C–C bond constructions, such as Minisci reactions, conjugate additions, and radical–radical couplings (*20*–*23*). Nevertheless, a general strategy for functionalization of C–H bonds via HAT–transition metal cross-coupling has yet to be achieved (*24*, *25*).

We recently questioned whether it would be possible to use a tertiary amine radical cation—generated via a photoredox-mediated single-

Merck Center for Catalysis at Princeton University, Princeton, NJ 08544, USA.
*These authors contributed equally to this work. †**Corresponding author. Email: dmacmill@princeton.edu**

electron transfer (SET) event (*23*, *26*–*28*)—to accomplish H-atom abstraction from a diverse range of substrates (Fig. 1). Given the electrophilic nature of amine radical cations, we proposed that such a catalytic strategy might allow the selective abstraction of hydridic, electron-rich C–H bonds in the presence of electron-deficient and neutral C–H bonds, which are abundant throughout organic molecules. We envisioned that the exploitation of polarity effects in the abstraction event would impart a high degree of kinetic selectivity into an otherwise unselective HAT process (*29*). Thereafter, we assumed the resulting radical intermediate might readily intersect with a Ni-catalyzed coupling cycle, thereby enabling C–C bond formation with a range of aryl electrophiles.

A detailed description of our proposed mechanistic cycle for the sp^3 C–H cross-coupling via photoredox HAT-nickel catalysis is outlined in Fig. 2. Initial excitation of the iridium(III) photocatalyst Ir[dF(CF_3)ppy]$_2$(dtbbpy)PF_6 [dF(CF_3)ppy = 2-(2,4-difluorophenyl)-5-(trifluoromethyl)pyridine; dtbbpy = 4,4'-di-*tert*-butyl-2,2'-bipyridine] (**1**) would produce the long-lived photoexcited state **2** (τ = 2.3 μs) (*30*). The *Ir(III) catalyst **2** is sufficiently oxidizing to undergo SET with a tertiary amine HAT catalyst (such as **3**), to generate Ir(II) **4** and amine radical cation **5** [$E_{1/2}^{red}$ (*Ir^{III}/Ir^{II}) = +1.21 V versus saturated calomel electrode (SCE) in CH_3CN; E_p (3-acetoxyquinuclidine) = +1.22 V versus SCE in CH_3CN] (*30*). As a central design element, we postulated that amine radical cation **5** would be sufficiently electron-deficient to engender a kinetically selective HAT event at the most electron-rich site of C–H nucleophile substrate **6**, thereby exclusively delivering radical intermediate **7**. At the same time, we hypothesized that this abstraction

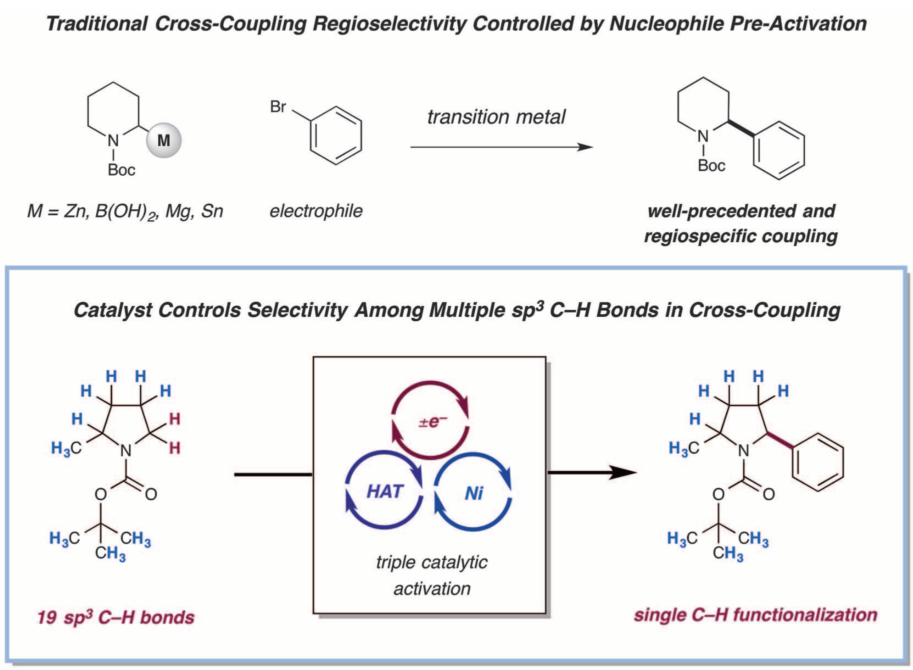


Fig. 1. Photoredox-mediated hydrogen atom transfer and nickel catalysis enables highly selective cross-coupling with sp^3 C–H bonds as latent nucleophiles.

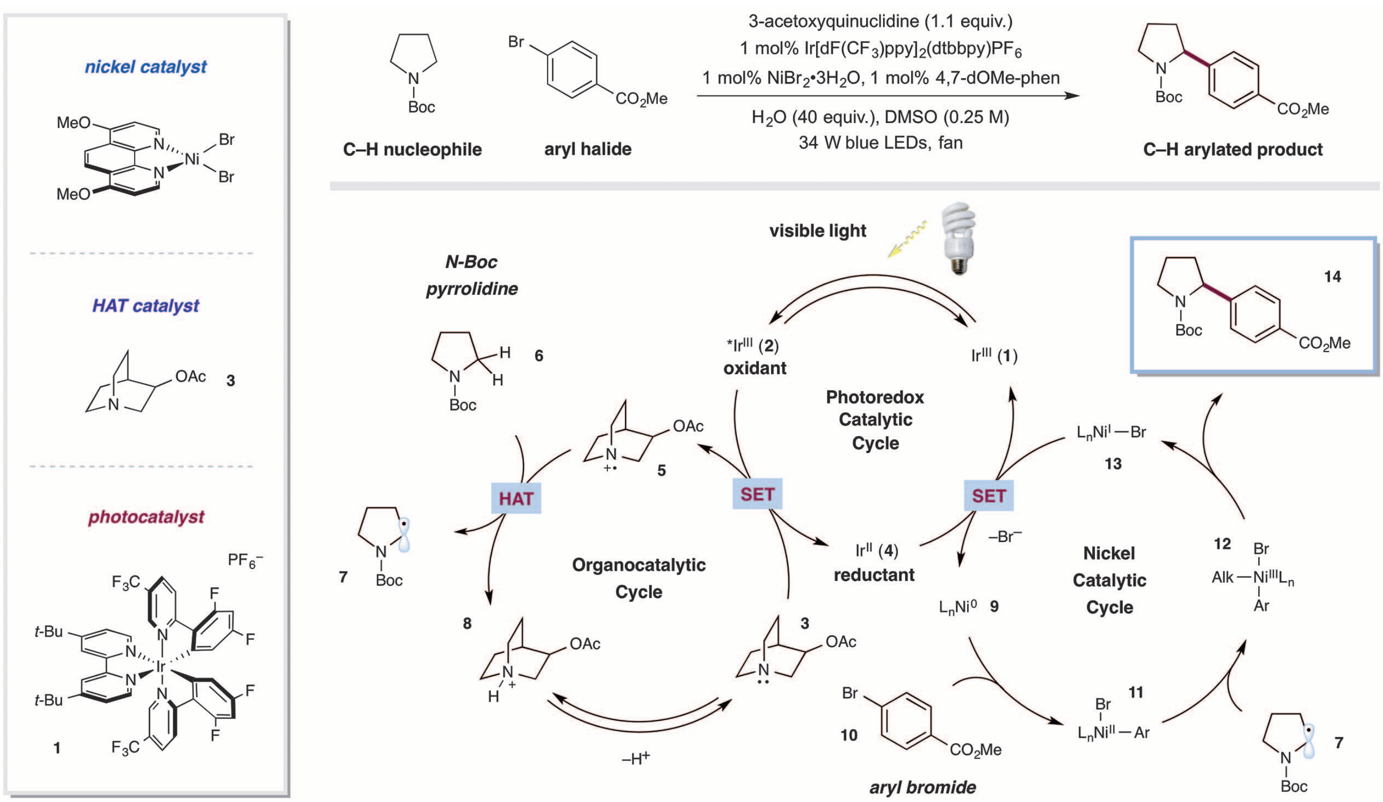


Fig. 2. Photoredox, HAT, and nickel-catalyzed cross-coupling: proposed mechanistic pathway and catalyst combination. Ac, acetyl; *t*-Bu, *tert*-butyl; Boc, *tert*-butoxycarbonyl; DMSO, dimethyl sulfoxide; LED, light-emitting diode; SET, single-electron transfer; HAT, hydrogen atom transfer.

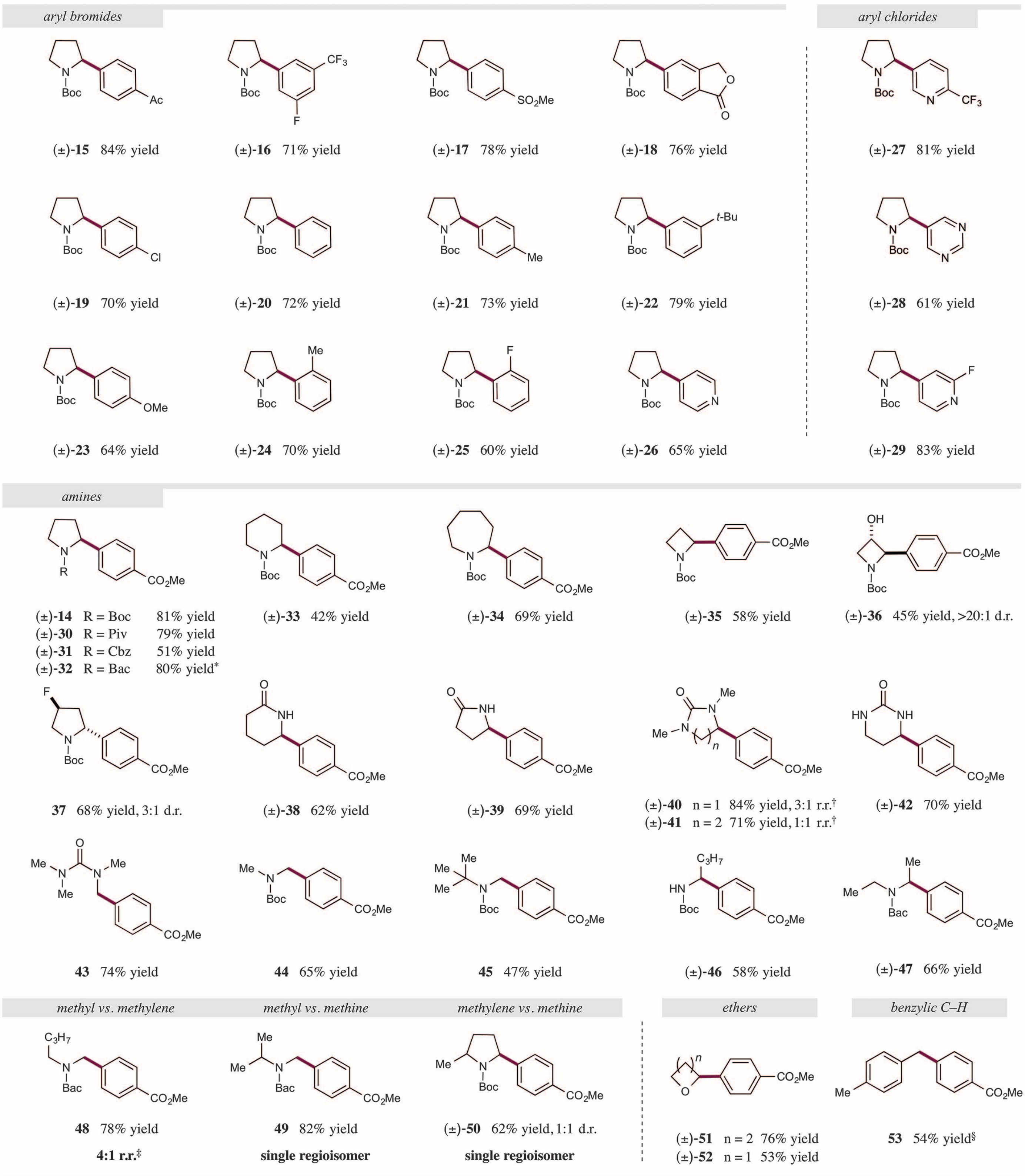


Fig. 3. Photoredox, HAT, and nickel-catalyzed cross-coupling: aryl halide and C–H nucleophile scope. All yields are isolated yields. Reaction conditions as in Fig. 2; see supplementary materials for experimental details. Ac, acetyl; *t*-Bu, *tert*-butyl; Boc, *tert*-butoxycarbonyl; Piv, pivalate; Cbz, benzyloxycarbonyl; Bac, *tert*-butylaminocarbonyl. *Reaction performed with 4-bromobenzotrifluoride to deliver *N*-Bac 2-(4-trifluoromethylphenyl)-pyrrolidine. †Minor regioisomer is arylated on Me position. ‡Minor regioisomer is arylated on α-amino methylene position. §Yield determined by ^{1}H–nuclear magnetic resonance.

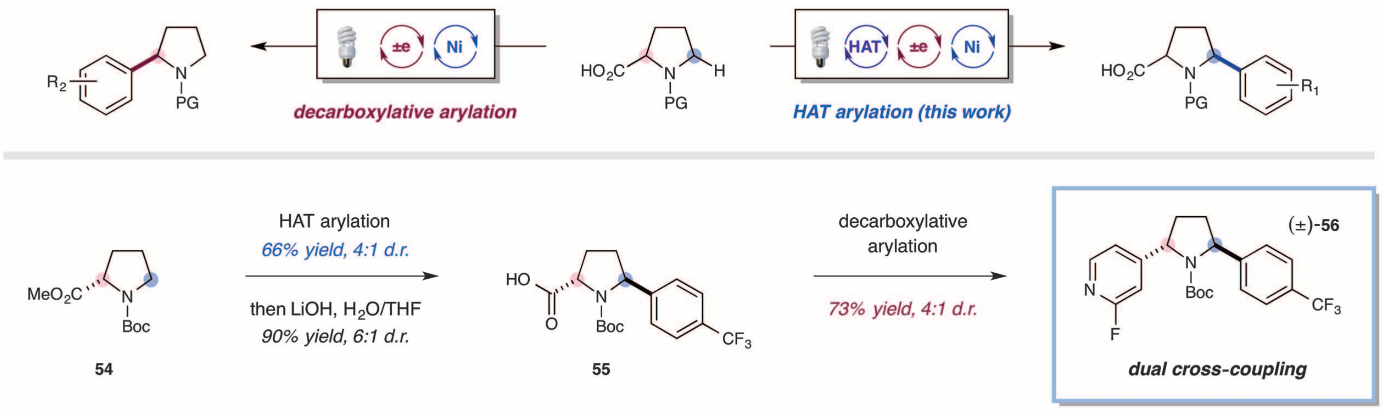


Fig. 4. Regioselective arylation: Using C–H arylation and decarboxylative arylation delivers differentially arylated pyrrolidine products. All yields are isolated yields. See supplementary materials for experimental details.

event should also be thermodynamically favorable considering the difference in the bond dissociation enthalpies (BDEs) of hydridic α- amino C–H bonds (α-amino C–H = 89 to 94 kcal/mol) and the resultant N–H bond of quinuclidinium cation [H–N$^+$ BDE (quinuclidine) = 100 kcal/mol] (*31*, *32*). Concurrent with this photoredox cycle, we assumed that our active Ni(0) species **9**—generated in situ via two SET reductions of (4,7-dOMe-phen)Ni(II)Br_2 (4,7-dOMe-phen = 4,7-dimethoxy-1,10-phenanthroline) by the iridium photocatalyst [$E_{1/2}^{red}$ (Ir^{III}/Ir^{II}) = –1.37 V versus SCE in CH_3CN; $E_{1/2}^{red}$ (Ni^{II}/Ni^{0}) = –1.2 V versus SCE in *N*,*N*-dimethylformamide] (*30*, *33*)—would undergo oxidative addition into the aryl halide electrophile **10**, forming the electrophilic Ni(II)-aryl intermediate **11**. This Ni(II) species would rapidly intercept radical **7** to generate a Ni(III)-aryl-alkyl complex **12**, which upon reductive elimination would forge the desired C–C bond to form Ni(I) complex **13** and benzylic amine **14**. Reduction of **13** by **4**, the Ir(II) state of the photocatalyst, would then reconstitute both Ni(0) catalyst **9** and Ir(III) catalyst **1**.

We began our investigations into the proposed photoredox-mediated HAT nickel cross-coupling by evaluating a broad range of photoredox catalysts, nickel-ligand systems, and quinuclidine analogs. Upon exposing *N*-Boc pyrrolidine and methyl 4-bromobenzoate to visible light [34 W blue light-emitting diodes (LEDs)] in the presence of iridium photocatalyst Ir[dF(CF_3)ppy]$_2$(dtbbpy)PF_6, $NiBr_2$•$3H_2O$, 4,7-dimethoxy-1,10-phenanthroline, and 3-acetoxyquinuclidine, we observed 81% yield of the desired α-amino C–C coupled product. Moreover, this product was the only detectable regioisomer formed, indicating that quinuclidine HAT catalyst **3** was selective for the most hydridic C–H bond available. Notably, using quinuclidine in lieu of 3-acetoxyquinuclidine resulted in diminished reactivity, indicating the necessity for an electron-withdrawing substituent. This substantial difference in reaction efficiency illustrates the capacity to tune the reactivity of the HAT catalyst via electronic modification of the substituent at the 3-position. It is important to note that under these reaction conditions, amine **3** serves as both the HAT catalyst and the base (*34*).

With the optimal conditions in hand, we next sought to examine the generality of this transformation by exploring the scope of the electrophilic aryl halide coupling partner. As outlined in Fig. 3, a wide variety of bromoarenes function efficiently in this HAT cross-coupling protocol. For example, electron-deficient aryl bromides containing ketones, trifluoromethyl groups, fluorines, sulfones, and esters were all effective arylating agents (**15** to **18**, 71 to 84% yield). Notably, 4-chlorobromobenzene gave chlorophenyl amine product **19** as the only observable arylation product in 70% yield, demonstrating that a high degree of chemoselectivity can be achieved in the oxidative addition step. The HAT arylation strategy is further effective for electron-neutral and electron-rich aryl bromides, as demonstrated by the installation of phenyl, tolyl, *t*-Bu-phenyl, and anisole groups (**20** to **23**, 64 to 79% yield). The presence of *ortho* methyl or fluorine substitution on the aryl halide was not problematic (**24** and **25**, 70 and 60% yield). With respect to heteroaromatic systems, pyridine rings were incorporated with good efficiency via the use of the corresponding heteroaryl bromide (**26**, 65% yield). Heteroaryl chlorides were also effective electrophiles in the transformation. For example, electron-deficient pyridines and pyrimidines deliver the benzylic amine products in good efficiency (**27** to **29**, 61 to 83% yield) (*35*). The collective one-step synthesis of the aryl pyrrolidine products **14** to **29** from simple *N*-Boc pyrrolidine clearly demonstrates that synthetic streamlining can be accomplished with this HAT cross-coupling technology (*36*).

We next explored the diversity of amino- and oxy-bearing C–H nucleophiles that could be used as substrates in this photoredox-mediated HAT nickel-catalyzed cross-coupling. As demonstrated in Fig. 3, many α-amino methyl- and methylene-containing substrates can be selectively arylated. For example, differentially *N*-substituted pyrrolidine substrates are effective in the transformation, including those bearing *tert*-butoxycarbonyl (Boc), benzyloxycarbonyl (Cbz), pivalate (Piv), and *tert*-butylaminocarbonyl (Bac) groups (**14**, **30** to **32**, 51 to 81% yield). Notably, the arylation of *N*-Boc pyrrolidine can be achieved on gram scale in a single batch, delivering 1.34 g of the 2-arylpyrrolidine product **14** (78% yield). Cyclic amines of various ring size are readily tolerated, with azetidine, piperidine, and azepane undergoing selective C–H arylation (**33** to **35**, 42 to 69% yield). Notably, ring systems that incorporate inductively withdrawing alcohols and fluorine substituents at the β-amino position do not unduly retard the C–H abstraction step [**36** and **37**, 45% yield, >20:1 diastereomeric ratio (d.r.). and 68% yield, 3:1 d.r.]. Moreover, lactams and ureas proved effective latent nucleophiles for this coupling, with both *N*-Me and *N*-H substrates providing the corresponding arylated products in good yield (**38** to **42**, 62 to 84% yield).

The transformation is not restricted to cyclic substrates, as a range of acyclic amines have been efficiently functionalized with this HAT arylation protocol. For example, primary α-amino C–H bonds in both *N*-Boc alkyl amines and ureas can be arylated in good yield (**43** to **45**, 47 to 74% yield). *N*-Boc butylamine, possessing a free N–H bond, undergoes selective α-arylation in 58% yield (**46**), leaving this latent functional handle available for further derivatization without the need for protection or deprotection steps. For acyclic dialkyl amines containing methyl and methylene C–H bonds, *N*-Bac–substituted amines delivered the α-arylated products in excellent yield (**47** to **49**, 66 to 82% yield), whereas the corresponding Boc systems provided diminished yet usable efficiencies (20 to 30% yield). We attribute this interesting reactivity difference to the diminished electron-withdrawing nature of the Bac group in comparison to Boc, resulting in an increased rate of hydrogen atom transfer to the electrophilic amine radical cation **5**.

When unsymmetrical amine substrates were exposed to this HAT protocol, some interesting regioselectivity patterns were discovered. For example, methyl C–H bonds undergo preferential coupling over methylene C–H bonds, as shown with *N*-Bac butylmethylamine [**48**, 78% yield, 4:1 regioisomeric ratio (r.r.)]. Furthermore, methyl and methylene C–H bonds react exclusively over methine C–H bonds, as demonstrated with *N*-Bac isopropylmethylamine and *N*-Boc 2-methylpyrrolidine, respectively (**49** and **50**, 82 and 62% yield,

1:1 d.r.). This strategy can also be applied to the HAT arylation of α-oxy C–H bonds. Tetrahydrofuran (THF) and oxetane both undergo α-oxy arylation in good efficiency (**51** and **52**, 76 and 53% yield). Finally, we have demonstrated that this C–H arylation protocol is effective for benzylic systems as *para*-xylene is arylated in 54% yield (**53**). Indeed, we expect that application of this strategy to a broad range of α-oxy, α-amino, and benzylic C–H–bearing substrates will demonstrate the general utility of this selective C–H arylation protocol.

Finally, the capacity to control the regioselectivity of the outlined HAT abstraction along with the opportunity to utilize C–H bonds as latent nucleophiles brings forward the possibility of enabling multiple native functionalizations to be conducted in sequence—a strategy that should allow the rapid construction of molecular complexity from a large variety of readily available organic feedstock chemicals. As one example, we postulated that *N*-Boc proline methyl ester (**54**) might be differentially arylated via (i) the photoredox-mediated HAT method presented in this work, followed by (ii) a photoredox-mediated Ni(II) decarboxylative arylation. As shown in Fig. 4, *N*-Boc proline methyl ester underwent selective arylation at the 5-methylene position using the HAT cross-coupling strategy described herein (66% yield, 4:1 d.r.). The observed regioselectivity is usefully complementary to that which would be expected with established methods for transition metal–catalyzed cross-coupling. Whereas many current strategies use basic conditions to selectively functionalize acidic hydrogens (as in enolate arylations), our developed HAT protocol targets hydridic hydrogen atoms, thereby providing access to fundamentally distinct product classes. Following the successful application of the C–H arylation outlined herein, the corresponding amino acid product **55** underwent decarboxylative coupling with 2-fluoro-4-bromopyridine at the 2-position, delivering the 2,5-diarylated pyrrolidine adduct in excellent yield (**56**, 73% yield, 4:1 d.r.). We have also demonstrated a HAT arylation followed by a nickel-catalyzed C–O coupling (*37*). *N*-Boc 3-hydroxyazetidine can be selectively arylated at the 2-position in 45% yield (**36**, Fig. 3), leaving the alcohol unreacted. The free alcohol can then be subsequently arylated with 4-bromo-2-methylpyridine to deliver the aryl ether product in 77% yield (see supplementary materials).

This HAT strategy represents a powerful demonstration of the versatility of using sp^3 C–H bonds as organometallic nucleophile equivalents and will likely find application in the realm of late-stage functionalization. We believe that this protocol will gain widespread use within the synthetic community as a complement to existing cross-coupling technologies.

REFERENCES AND NOTES

1. M. Beller, C. Bolm, *Transition Metals for Organic Synthesis, Vol. 1* (Wiley-VCH, Weinheim, 2004).
2. A. De Meijere, F. Diederich, *Metal-Catalyzed Cross-Coupling Reactions, Vol. 1* (Wiley-VCH, Weinheim, 2004).
3. A. de Meijere, F. E. Meyer, *Angew. Chem. Int. Ed. Engl.* **33**, 2379–2411 (1995).
4. L. J. Gooßen, G. Deng, L. M. Levy, *Science* **313**, 662–664 (2006).
5. N. Rodríguez, L. J. Goossen, *Chem. Soc. Rev.* **40**, 5030–5048 (2011).
6. L. L. Anka-Lufford, M. R. Prinsell, D. J. Weix, *J. Org. Chem.* **77**, 9989–10000 (2012).
7. J. C. Lo, J. Gui, Y. Yabe, C.-M. Pan, P. S. Baran, *Nature* **516**, 343–348 (2014).
8. Z. Zuo *et al.*, *Science* **345**, 437–440 (2014).
9. V. G. Zaitsev, D. Shabashov, O. Daugulis, *J. Am. Chem. Soc.* **127**, 13154–13155 (2005).
10. S. J. Pastine, D. V. Gribkov, D. Sames, *J. Am. Chem. Soc.* **128**, 14220–14221 (2006).
11. J. He *et al.*, *Science* **343**, 1216–1220 (2014).
12. F.-L. Zhang, K. Hong, T.-J. Li, H. Park, J.-Q. Yu, *Science* **351**, 252–256 (2016).
13. J. J. Topczewski, P. J. Cabrera, N. I. Saper, M. S. Sanford, *Nature* **531**, 220–224 (2016).
14. M. Lee, M. S. Sanford, *J. Am. Chem. Soc.* **137**, 12796–12799 (2015).
15. J. C. Lewis, P. S. Coelho, F. H. Arnold, *Chem. Soc. Rev.* **40**, 2003–2021 (2011).
16. J. J. Warren, T. A. Tronic, J. M. Mayer, *Chem. Rev.* **110**, 6961–7001 (2010).
17. J. M. Mayer, *Acc. Chem. Res.* **44**, 36–46 (2011).
18. M. S. Chen, M. C. White, *Science* **327**, 566–571 (2010).
19. T. Newhouse, P. S. Baran, *Angew. Chem. Int. Ed.* **50**, 3362–3374 (2011).
20. K. Qvortrup, D. A. Rankic, D. W. C. MacMillan, *J. Am. Chem. Soc.* **136**, 626–629 (2014).
21. J. D. Cuthbertson, D. W. C. MacMillan, *Nature* **519**, 74–77 (2015).
22. J. Jin, D. W. C. MacMillan, *Nature* **525**, 87–90 (2015).
23. J. L. Jeffrey, J. A. Terrett, D. W. C. MacMillan, *Science* **349**, 1532–1536 (2015).
24. D. Liu, C. Liu, H. Li, A. Lei, *Angew. Chem. Int. Ed.* **52**, 4453–4456 (2013).
25. D. Liu *et al.*, *Org. Lett.* **17**, 998–1001 (2015).
26. J. M. R. Narayanam, C. R. J. Stephenson, *Chem. Soc. Rev.* **40**, 102–113 (2011).
27. C. K. Prier, D. A. Rankic, D. W. C. MacMillan, *Chem. Rev.* **113**, 5322–5363 (2013).
28. D. M. Schultz, T. P. Yoon, *Science* **343**, 1239176 (2014).
29. B. P. Roberts, *Chem. Soc. Rev.* **28**, 25–35 (1999).
30. M. S. Lowry *et al.*, *Chem. Mater.* **17**, 5712–5719 (2005).
31. Y.-R. Luo, *Handbook of Bond Dissociation Energies in Organic Compounds* (CRC Press, Boca Raton, FL, 2003).
32. W.-Z. Liu, F. G. Bordwell, *J. Org. Chem.* **61**, 4778–4783 (1996).
33. M. Durandetti, M. Devaud, J. Perichon, *New J. Chem.* **20**, 659 (1996).
34. In related studies, we have demonstrated that HAT processes can be achieved with catalytic quantities of 3-acetoxyquinuclidine in the presence of base.
35. For heteroaryl halides, no product was observed in the absence of nickel catalyst, supporting the cross-coupling mechanism outlined in Fig. 2. See supplementary materials for control reactions.
36. K. R. Campos, A. Klapars, J. H. Waldman, P. G. Dormer, C. Y. Chen, *J. Am. Chem. Soc.* **128**, 3538–3539 (2006).
37. J. A. Terrett, J. D. Cuthbertson, V. W. Shurtleff, D. W. C. MacMillan, *Nature* **524**, 330–334 (2015).

ACKNOWLEDGMENTS

We are grateful for financial support provided by the NIH National Institute of General Medical Sciences (R01 GM078201-05) and gifts from Merck and AbbVie. V.W.S. and J.A.T. thank Bristol-Myers Squibb for graduate fellowships. J.D.C. thanks Marie Curie Actions for an International Outgoing Fellowship.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1304/suppl/DC1
Tables S1 and S2
NMR Spectra
References (*38*–*50*)

10 March 2016; accepted 14 April 2016
Published online 28 April 2016
10.1126/science.aaf6635

GLASS TRANSITION

Fifth-order susceptibility unveils growth of thermodynamic amorphous order in glass-formers

S. Albert,[1] Th. Bauer,[2]* M. Michl,[2] G. Biroli,[3,4] J.-P. Bouchaud,[5] A. Loidl,[2] P. Lunkenheimer,[2] R. Tourbot,[1] C. Wiertel-Gasquet,[1] F. Ladieu[1]†

Glasses are ubiquitous in daily life and technology. However, the microscopic mechanisms generating this state of matter remain subject to debate: Glasses are considered either as merely hyperviscous liquids or as resulting from a genuine thermodynamic phase transition toward a rigid state. We show that third- and fifth-order susceptibilities provide a definite answer to this long-standing controversy. Performing the corresponding high-precision nonlinear dielectric experiments for supercooled glycerol and propylene carbonate, we find strong support for theories based on thermodynamic amorphous order. Moreover, when lowering temperature, we find that the growing transient domains are compact—that is, their fractal dimension $d_f = 3$. The glass transition may thus represent a class of critical phenomena different from canonical second-order phase transitions for which $d_f < 3$.

The glassy state of matter, despite its omnipresence in nature and technology (*1*), continues to be one of the most puzzling riddles in condensed-matter physics (*1*, *2*): For all practical purposes, glasses are rigid like crystals, but they lack any long-range order. Some theories describe glasses as kinetically constrained liquids (*3*), becoming so viscous below the glass transition that they seem effectively rigid. By contrast, other theories (*4*, *5*) are built on the existence of an underlying thermodynamic phase transition to a state where the molecules are frozen in well-defined yet disordered positions. This so-called "amorphous order" cannot be revealed by canonical static correlation functions, but rather by new kinds of correlations [i.e., point-to-set correlations or other measures of local order (*6*, *7*)] that have been detected in recent numerical simulations (*7*–*9*). In these theories, thermodynamic correlations lock together

the fluctuations and response of the molecules, which collectively rearrange over some length scale ℓ, ultimately leading to rigidity. In this thermodynamic scenario, ℓ is proportional to a power of $\ln(\tau_\alpha/\tau_0)$, where τ_α is the structural relaxation time and τ_0 is the microscopic time scale, generally smaller than 1 ps (4, 5). Because equilibrium measurements require a time longer than τ_α, they cannot be performed in the range where ℓ is very large, which would require exponentially long times. This limitation is essentially why the true nature of glasses is still a matter of intense debate.

Here, we propose a strategy to unveil the existence of a thermodynamic length ℓ that grows upon cooling. Instead of only varying the temperature T, we also vary the nonlinear order k of the response of supercooled liquids. This is motivated by a general, although rarely considered (10), property of critical points: At a second-order critical temperature T_c, the linear susceptibility χ_1 associated with the order parameter is not the only diverging response. As a function of temperature, all the higher-order responses χ_{2m+1} ($m \geq 1$) diverge even faster than χ_1 itself. This comes from the fact that the divergences of all the χ_{2m+1} have the same origin—namely, the divergence of the length ℓ. By using the appropriate scaling theory, it can be shown that the larger the value of m, the stronger the divergence in temperature.

[1]SPEC, CEA, CNRS, Université Paris-Saclay, CEA Saclay Bat 772, 91191 Gif-sur-Yvette Cedex, France. [2]Experimental Physics V, Center for Electronic Correlations and Magnetism, University of Augsburg, 86159 Augsburg, Germany. [3]IPhT, CEA, CNRS, Université Paris-Saclay, CEA Saclay Bat 774, 91191 Gif-sur-Yvette Cedex, France. [4]LPS, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France. [5]Capital Fund Management, 23 rue de l'Université, 75007 Paris, France. *Present address: Institute for Machine Tools and Industrial Management, Technical University of Munich, 85748 Garching, Germany. †**Corresponding author. Email: francois.ladieu@cea.fr**

As theoretically shown below, transposing this idea to glasses requires taking into account that the putative "amorphous" or hidden order in supercooled liquids (7, 11) is not reflected in χ_1 itself, but only in higher-order response functions χ_{2m+1} ($m \geq 1$). This idea is indeed supported by previous measurements and analyses of the third-order susceptibility χ_3 (12–16). We report results on the fifth-order susceptibility $\chi_5(T)$ and compare them to $\chi_3(T)$ in two canonical glass forming-liquids, glycerol and propylene carbonate. If critical phenomena really play a key role for the glass transition, χ_5 should increase much faster than χ_3 as the liquid becomes more viscous.

This scenario can be understood by means of a theoretical argument based on previous work (17) and further detailed in (18). Suppose that $N_{corr} = (\ell/a)^{d_f}$ molecules are amorphously ordered over the length scale ℓ, where a is the molecular size and d_f is the fractal dimension of the ordered clusters. This implies that their dipoles, oriented in apparently random positions, are essentially locked together during a time τ_α. We expect that in the presence of an external electric field E oscillating at frequency $\omega \geq \tau_\alpha^{-1}$, the dipolar degrees of freedom of these molecules contribute to the polarization per unit volume as

$$p = \mu_{dip} \frac{\sqrt{(\ell/a)^{d_f}}}{(\ell/a)^d} F\left[\frac{\mu_{dip} E \sqrt{(\ell/a)^{d_f}}}{kT}\right] \qquad (1)$$

where μ_{dip} is an elementary dipole moment, F is a scaling function such that $F(-x) = -F(x)$, and $d = 3$ is the dimension of space. This states that randomly locked dipoles have an overall moment $\sim \sqrt{N_{corr}}$ and that we should compare the energy of this "super-dipole" in a field to the thermal energy. Equation 1 is motivated by general arguments involving multipoint correlation functions through which d_f can be given a precise meaning (18) and is fully justified when ℓ diverges, in particular in the vicinity of a critical point such as the mode-coupling transition or the spin-glass transition. In the latter case, Eq. 1 is in fact equivalent to the scaling arguments of (19), provided one performs suitable mapping between the magnetic formalism of (19) and ours.

Expanding Eq. 1 in powers of E, we find the "glassy" contribution to p:

$$\frac{p}{\mu_{dip}} = F'(0)\left(\frac{\ell}{a}\right)^{d_f-d}\left(\frac{\mu_{dip}E}{kT}\right) + \frac{1}{3!}F^{(3)}(0)\left(\frac{\ell}{a}\right)^{2d_f-d}\left(\frac{\mu_{dip}E}{kT}\right)^3 + \frac{1}{5!}F^{(5)}(0)\left(\frac{\ell}{a}\right)^{3d_f-d}\left(\frac{\mu_{dip}E}{kT}\right)^5 + \ldots \qquad (2)$$

Because d_f must be less than or equal to d, we find that the first term, contributing to the usual linear dielectric constant $\chi_1(\omega)$, cannot grow as ℓ increases. This simple theoretical argument explains why we do not expect spatial glassy correlations to show up in $\chi_1(\omega)$. The second term, contributing to the third-order dielectric constant, does grow with ℓ, provided $d_f > d/2$. Although $d_f < d$ close to a standard second-order critical point (20) such as the spin-glass transition, several theories suggest (4, 5, 21, 22) that ordered domains are compact ($d_f = d$), in which case $(\ell/a)^{2d_f-d} = (\ell/a)^d = N_{corr}$, as assumed in our previous work (17, 23). The third term of Eq. 2 reveals that the fifth-order susceptibility $\chi_5(\omega)$ should diverge as ℓ^{3d_f-d}. Therefore, the joint measurement of $\chi_3(\omega)$ and $\chi_5(\omega)$ provides a direct way to estimate d_f

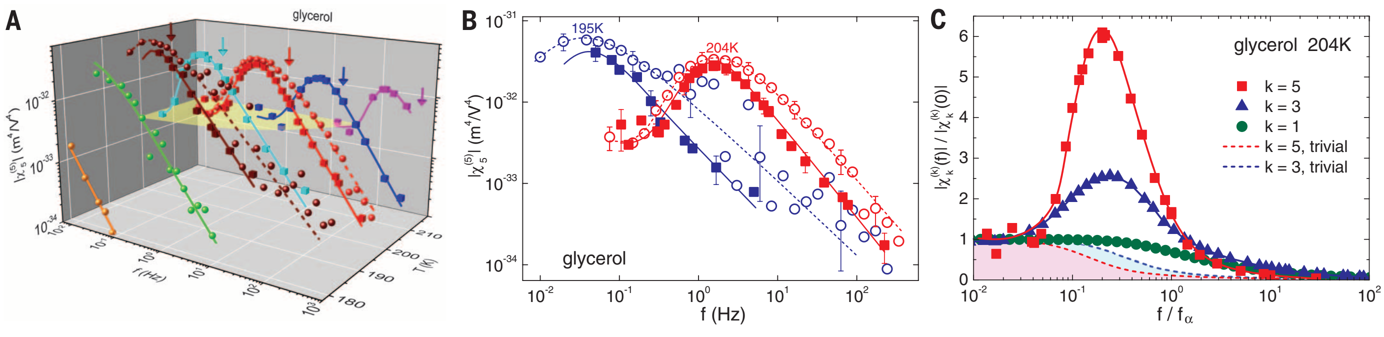


Fig. 1. Modulus of the fifth-order susceptibility in supercooled glycerol measured with two independent setups. (**A**) The susceptibilities $\chi_5^{(5)}$ reported here are obtained directly (18) by monitoring the response of the sample at 5ω when applying an electric field E at angular frequency ω. Two independent setups were used, designed either to maximize the field amplitude (Augsburg setup, spheres) or to optimize the sensitivity (Saclay setup, cubes). Lines are guides to the eye. Errors are on the order of the scatter of neighboring data points around the lines. Both setups yield consistent results. For a given temperature T, $|\chi_5^{(5)}|$ has a humped shape, with a maximum occurring at the frequency $f_{peak} \approx 0.22 f_\alpha$ where f_α is the relaxation frequency indicated by a colored arrow for each temperature. When decreasing T, the height of the hump increases strongly. The yellow plane emphasizes the fact that, for a given T, $\chi_5^{(5)}$ is constant for $f/f_\alpha \leq 0.05$. (**B**) Projection onto the susceptibility-frequency plane of the data in Fig. 1A at 204 K and at 195 K. The agreement around and below the peak is remarkable at 204 K (see text). The relative evolution of the height of the peak is reasonably similar between 204 K and 195 K for the two setups (see Fig. 3A). (**C**) Comparison of the fifth-order, cubic, and linear susceptibilities of glycerol [the latter is notated $\chi_1^{(1)}$ for convenience (18)]. Symbols, with lines to guide the eye, are Saclay data at 204 K; the error bars are on the order of the size of the symbols for $k = 5$ [except at the lowest frequencies (18)] and smaller for $k = 3$ and 1. The higher the order k, the stronger the hump of $|\chi_k^{(k)}|$; this is a key result supporting the amorphous-order scenario. The dashed lines, emphasized by colored areas, correspond to the trivial response of an ideal gas of dipoles without amorphous order. In this case, $|\chi_k^{(k)}|$ decreases monotonously in frequency for any value of k. The higher k, the stronger the difference between the measured and trivial susceptibility.

experimentally through the following relation:

$$|\chi_5| \propto |\chi_3|^{\mu(d_f)};\ \mu(d_f) = \frac{3d_f - d}{2d_f - d};\ d_f(\mu) = d\,\frac{\mu - 1}{2\mu - 3} \quad (3)$$

where the exponent $\mu(d_f)$ is equal to 2 when the dynamically correlated regions are compact ($d_f = d$), and is higher otherwise. We predict two key results that can be obtained from χ_5 and χ_3 susceptibility measurements. First, if amorphous order increases approaching the transition, the frequency dependence should be more anomalous [i.e., more humped (18)] for $\chi_5(\omega)$ than for $\chi_3(\omega)$. Second, the growth of χ_5 should be much stronger than that of χ_3 when lowering the temperature, following $\chi_5 \sim \chi_3^2$ if we assume compact amorphous domains. Our work provides experimental evidence that these predictions indeed hold and suggests that the glass transition represents a new type of critical phenomenon with growing length and time scales but with $d_f = d$, in contrast to the spin-glass transition that instead displays (19) canonical critical behavior with $d_f \approx 2.35$.

We measured $\chi_5(\omega)$ in glycerol and propylene carbonate by applying a field of amplitude E and frequency $f = \omega/(2\pi)$ (18). The fifth-order response is $\propto \chi_5 E^5$ and is orders of magnitude smaller than the cubic and linear ones, given by $\propto \chi_3 E^3$ and $\propto \chi_1 E$, respectively. We avoided any contributions of χ_3 and of χ_1 by measuring the signal at 5ω, which only contains the component $\chi_5^{(5)}$ of the fifth-order susceptibility (18). We measured $\chi_5^{(5)}$ with two independent setups because of the very small amplitude, optimized along complementary strategies. One setup (in Augsburg) was designed to achieve the highest possible field (reaching 78 MV/m). We optimized sensitivity with a differential technique using two samples of different thicknesses in the other setup (Saclay; see fig. S1), which required lower fields (up to 26 MV/m).

We obtained the values of $|\chi_5^{(5)}(\omega)|$ for glycerol at various frequencies and temperatures by using the two aforementioned techniques (Fig. 1A). A clear peak arises for a given T in $|\chi_5^{(5)}(\omega)|$ for a frequency $f_{peak} \approx 0.22 f_\alpha$ where the α-relaxation frequency f_α, defined by the peak of the out-of-phase linear susceptibility, is indicated by arrows in Fig. 1A. Even though the data were determined by two independent setups, the overall agreement is remarkable (Fig. 1B). The most accurate comparison is possible at 204 K, where f_{peak} is well inside the frequency range accessible by the two setups. The two spectra at 204 K coincide on the low-frequency side of the peak (18). On the other side of the peak, a discrepancy between the two sets of data progressively increases with frequency, reaching a constant factor of 4 at the highest frequencies (Fig. 1B). Apart from the value of the electric field, the main difference between the two experiments is the number of applied field cycles n. The Saclay setup measured the stationary responses ($n \to \infty$), whereas n remained finite in the Augsburg setup [similarly to (24)], ranging from $n = 2$ at the lowest frequencies to $n \propto f$ at the highest frequencies. The two setups give the same results for $\chi_5^{(5)}$ because at sufficiently low values of f/f_α, the response of

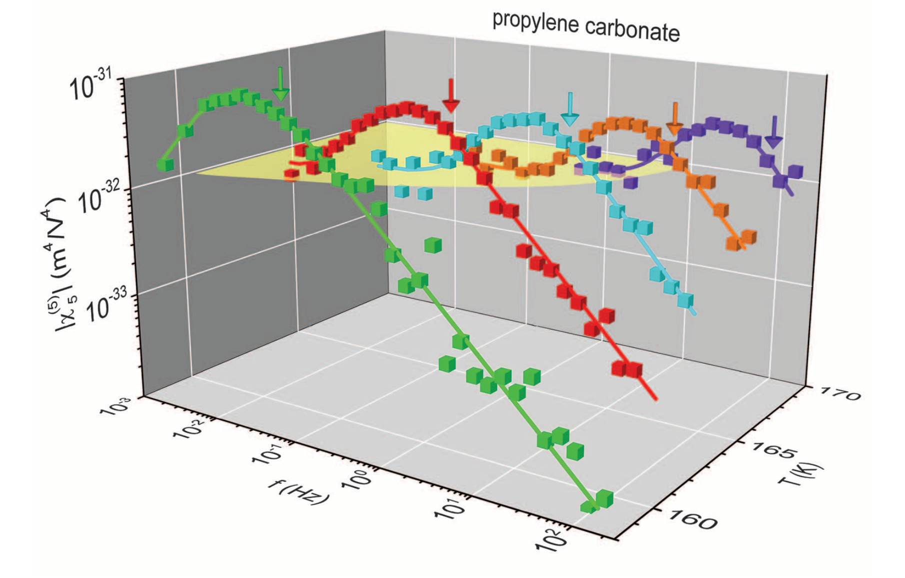


Fig. 2. Modulus of the fifth-order susceptibility in supercooled propylene carbonate. The experimental data (symbols) were obtained with the Augsburg setup. The presentation of the graph is analogous to Fig. 1A to emphasize the similarity of the behavior of $|\chi_5^{(5)}|$ in propylene carbonate and in glycerol, even though these two liquids have different fragilities and different types of intermolecular interactions (van der Waals bonding versus hydrogen bonding).

the supercooled liquid is likely to instantaneously follow the field. By contrast, at higher frequencies $\omega \geq \tau_\alpha^{-1}$, the finite cycle number may play a role, making a quantitative treatment of this effect difficult (18). Our further analysis relies on the behavior of the peaks of $\chi_5^{(5)}$, and more precisely on their relative evolution with temperature, which reasonably agrees in the two setups (see below).

The qualitative features of $|\chi_5^{(5)}(\omega)|$ (Fig. 1, A and B) are reminiscent of those of the third-harmonic cubic susceptibility $|\chi_3^{(3)}(\omega)|$ (12, 13). Both quantities exhibit a humped shape, with a peak located at the same frequency $f_{peak} \approx 0.22 f_\alpha$, as well as a strong increase of the height of the peak as the temperature is decreased. These two distinctive features are important because they are specific signatures of glassy dynamical correlations (17), in contrast to trivial systems without correlations (25). In this case, the modulus of all higher-order nonlinear susceptibilities monotonously decreases with frequency (18, 25).

To quantitatively compare the frequency dependency of the susceptibilities $\chi_k^{(k)}$ of order k, we plotted $|\chi_k^{(k)}(f/f_\alpha)/\chi_k^{(k)}(0)|$ of glycerol for k = 5, 3, and 1 ($\chi_1^{(1)}$ is the linear susceptibility noted χ_1 above) (Fig. 1C). The peak amplitude for $k = 5$ is strongly enhanced relative to $k = 3$—that is, the higher the nonlinear order k, the more anomalous the frequency dependence (Fig. 1C and figs. S2 and S3). This behavior is a decisive result and is fully consistent with our scaling analysis. For archetypical glass-formers, we can always fit the linear susceptibility by assuming a sum of Debye relaxations where $\chi_{1,\mathrm{Debye}} \propto 1/(1 - i\omega\tau)$. We do this by choosing a suitable distribution $G(\tau)$ of relaxation times τ (26) caused by dynamical heterogeneities. Because the trivial response discussed above also obeys $\chi_{1,\mathrm{trivial}} \propto 1/(1 - i\omega\tau)$, we have used (18, 25) the same distribution $G(\tau)$ to calculate the trivial responses $\chi_{k,\mathrm{trivial}}^{(k)}$ for k = 3 and 5. For a given $k > 1$, a large difference exists between the experimental spectrum of $|\chi_k^{(k)}(f/f_\alpha)/\chi_k^{(k)}(0)|$ and its trivial counterpart (Fig. 1C), which we can ascribe to correlation-induced effects. For $k = 1$, the experimental data agree with the trivial response [convoluted with $G(\tau)$], consistent with the theoretical arguments stating that glassy correlations do not change the linear response (17). For k = 3 and 5, the difference to the trivial response increases, being much more important for $k = 5$, where it exceeds one order of magnitude. This quantitatively supports the scaling prediction obtained assuming that collective effects due to the growth of amorphous order play a key role in supercooled liquids.

We measured $|\chi_5^{(5)}(\omega)|$ at five different temperatures for propylene carbonate (Fig. 2). Propylene carbonate differs from glycerol in that its fragility (27, 28) $m \propto [\partial \log(\tau_\alpha)/\partial(1/T)]_{T_g}$ (where T_g is the glass transition temperature) is twice as large and it has van der Waals bonding, in contrast to hydrogen bonding. Despite these differences, the anomalous hump-like features of glycerol (Fig. 1A) are also observed in propylene carbonate (Fig. 2). We expect this behavior from our scaling framework, which relies on the predominant role of collective dynamical effects in supercooled liquids. The presence of similar anomalous features in two very different glass-formers suggests that they only weakly depend on the specific microscopic properties of the material.

To elicit the temperature dependence of collective effects, we introduced dimensionless quantities related to $\chi_3^{(3)}$ and $\chi_5^{(5)}$:

$$X_3^{(3)} \equiv \frac{k_B T}{\varepsilon_0 \Delta\chi_1^2 a^3}\chi_3^{(3)}, \qquad X_5^{(5)} \equiv \frac{(k_B T)^2}{\varepsilon_0^2 \Delta\chi_1^3 a^6}\chi_5^{(5)} \quad (4)$$

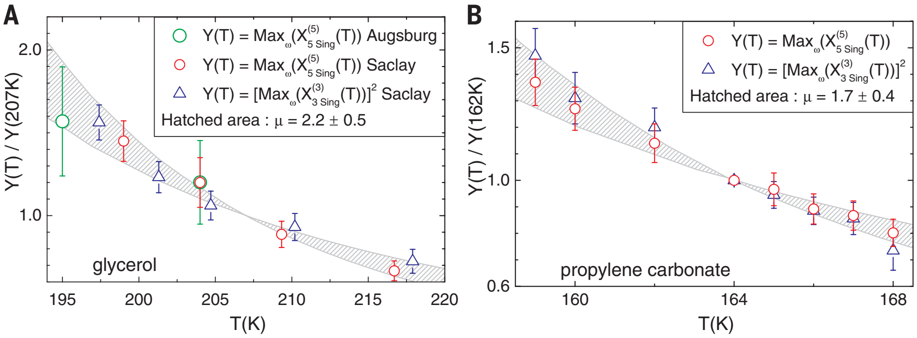


Fig. 3. Comparison of the temperature dependence of the singular part of the fifth-order and cubic dimensionless susceptibilities at f_{peak}. (**A**) For glycerol, the singular part of $|X_k^{(k)}(f_{peak})|$ for $k = 3$ and 5 is normalized to 1 at 207 K. The value of the exponent μ is then determined by comparing $|X^{(5)}_{5,sing}(f_{peak})|$ to $|X^{(3)}_{3,sing}(f_{peak})|^{\mu}$; the symbols for $k = 3$ correspond to $\mu = 2$, and the hatched area shows the interval corresponding to the error bar given for μ (*18*). The two Augsburg data points for $X^{(5)}_5$ have been added on the graph by scaling to the Saclay point at 204 K; the Augsburg point at 195 K is reasonably well within the hatched area, which shows that the relative evolution of $X^{(5)}_5$ with temperature is consistent in the two setups. (**B**) Same display as in (A), but for propylene carbonate with $T = 164$ K as the normalization temperature and the symbols for $k = 3$ corresponding to $\mu = 2$.

where ε_0 is the permittivity of free space, $\Delta\chi_1 = \chi_1(0) - \chi_1(\infty)$ is the dielectric strength, a^3 is the molecular volume, and k_B is the Boltzmann constant. The main advantage of these dimensionless nonlinear susceptibilities is that in the trivial case of an ideal gas of dipoles, both $X^{(3)}_{3,trivial}$ and $X^{(5)}_{5,trivial}$ are independent of temperature when plotted versus scaled frequency (*18*, *25*). Hence, we ascribe their experimental variation to the nontrivial dynamical correlations in the supercooled liquid (*17*, *23*). This interpretation is strongly supported by previous findings (*12*–*14*, *23*) where the temperature dependence of $|X^{(3)}_3|$ was studied at various values of f/f_α. Close to and above its peak frequency, $|X^{(3)}_3|$ was found to strongly vary in temperature, contrary to the low-frequency plateau region ($f/f_\alpha \leq 0.05$) where $|X^{(3)}_3|$ no longer depends on temperature. This low-frequency region corresponds to time scales much longer than τ_α where the liquid flow destroys glassy correlations, making each molecule effectively independent of others and yielding a dielectric response close to the aforementioned trivial case. This is why, to determine the temperature evolution of the glassy dynamical correlations, we focused on the region of the peak of $|\chi^{(5)}_5|$. For each of the two liquids, this peak appears at the very same frequency f_{peak} as in $|\chi^{(3)}_3|$.

We expect the nonlinear susceptibilities to contain a trivial contribution that would exist even for independent dipoles, as well as a "singular" contribution (i.e., diverging with ℓ) as given in Eq. 2. We thus write

$$X^{(3)}_{3,sing} \equiv X^{(3)}_3 - X^{(3)}_{3,trivial}, \qquad X^{(5)}_{5,sing} \equiv X^{(5)}_5 - X^{(5)}_{5,trivial} \tag{5}$$

Here, the trivial contributions are calculated by assuming a set of independent Debye dipoles convoluted with the aforementioned distribution $G(\tau)$ of relaxation times (*18*). We compared the temperature evolution of $|X^{(5)}_{5,sing}[f_{peak}(T)]|$ and that of $|X^{(3)}_{3,sing}[f_{peak}(T)]|^{\mu}$ (Fig. 3) to derive the value of the exponent μ, from which we deduce the fractal dimension d_f of the dynamically correlated regions by using Eq. 3. In both glycerol and propylene carbonate, the value $\mu = 2$, corresponding to compact domains of dimension $d_f = 3$, is found to be consistent with experiments (triangles in Fig. 3). By fitting the T dependence of $|X^{(3)}_{3,sing}[f_{peak}(T)]|$ with a smooth function (*18*), we found the hatched area corresponding to $\mu = 2.2 \pm 0.5$ in glycerol and $\mu = 1.7 \pm 0.4$ in propylene carbonate (Fig. 3). The fact that, within experimental uncertainty, a value of $\mu \approx 2$ is common to each of the two liquids supports a picture of amorphous compact domains mostly independent of differences at the molecular level and validates the correlation length scale for our scaling analysis. Considering that the temperature interval in Fig. 3B is smaller by a factor of 2, we note that the critical behavior in propylene carbonate is stronger than in glycerol (Fig. 3A). This suggests that the larger the fragility, the stronger the temperature dependence of the thermodynamic length ℓ. This is easily understood in the scenario of (*4*), where the critical point is the Vogel-Fulcher temperature T_0: In this case, equilibrium measurements can be made closer to the critical point for more fragile liquids, because the larger the fragility, the smaller the difference between T_g and T_0.

Our experimental results are therefore consistent with the general predictions of theories such as the random first-order transition or frustration-limited domains (*4*, *5*), where the physical mechanism driving the glass transition is of thermodynamic origin and where some nontrivial (albeit random) long-range correlations build up between molecules. Only in this case (*18*) can one have N_{corr} dipolar degrees of freedom collectively responding over some length scale ℓ and over time scales on the order of τ_α. If instead the glass transition is regarded as a purely dynamical phenomenon, there would not be any anomalous increase of the normalized peak value of the higher-order susceptibilities at all (*18*). Our results therefore severely challenge theories advocating against any thermodynamic signature and favoring purely dynamic scenarios. Moreover, from a comparison of the higher-order susceptibilities, our results are consistent with $\chi_5 \propto \chi_3^2$. This constitutes evidence for compact amorphously ordered domains (i.e., $d_f = d$) pointing toward a nonstandard nature of the glass transition, in contrast to canonical second-order phase transitions for which $d_f < d$.

REFERENCES AND NOTES

1. M. D. Ediger, P. Harrowell, *J. Chem. Phys.* **137**, 080901 (2012).
2. L. Berthier, G. Biroli, *Rev. Mod. Phys.* **83**, 587–645 (2011).
3. D. Chandler, J. P. Garrahan, *Annu. Rev. Phys. Chem.* **61**, 191–217 (2010).
4. M. Dzero, J. Schmalian, P. G. Wolynes, in *Structural Glasses and Supercooled Liquids: Theory, Experiment, and Applications*, P. G. Wolynes, V. Lubchenko, Eds. (Wiley, 2012), pp. 193–222.
5. G. Tarjus, S. A. Kivelson, Z. Nussinov, P. Viot, *J. Phys. Condens. Matter* **17**, R1143–R1182 (2005).
6. S. Mossa, G. Tarjus, *J. Chem. Phys.* **119**, 8069 (2003).
7. G. Biroli, J.-P. Bouchaud, A. Cavagna, T. S. Grigera, P. Verrocchio, *Nat. Phys.* **4**, 771–775 (2008).
8. R. Pinney, T. B. Liverpool, C. P. Royall, *J. Chem. Phys.* **143**, 244507 (2015).
9. L. Berthier, P. Charbonneau, S. Yaida, *J. Chem. Phys.* **144**, 024501 (2016).
10. Y. Kimura, S. Hara, R. Hayakawa, *Phys. Rev. E* **62**, R5907–R5910 (2000).
11. J.-P. Bouchaud, G. Biroli, *J. Chem. Phys.* **121**, 7347–7354 (2004).
12. C. Crauste-Thibierge *et al.*, *Phys. Rev. Lett.* **104**, 165703 (2010).
13. T. Bauer, P. Lunkenheimer, A. Loidl, *Phys. Rev. Lett.* **111**, 225702 (2013).
14. C. Brun, F. Ladieu, D. L'Hôte, G. Biroli, J.-P. Bouchaud, *Phys. Rev. Lett.* **109**, 175702 (2012).
15. U. Buchenau, R. Zorn, M. A. Ramos, *Phys. Rev. E* **90**, 042312 (2014).
16. R. Casalini, D. Fragiadakis, C. M. Roland, *J. Chem. Phys.* **142**, 064504 (2015).
17. J.-P. Bouchaud, G. Biroli, *Phys. Rev. B* **72**, 064204 (2005).
18. See supplementary materials on *Science* Online.
19. L. P. Lévy, *Phys. Rev. B* **38**, 4963–4973 (1988).
20. A. Coniglio, A. Fierro, in *Encyclopedia of Complexity and Systems Science*, R. A. Meyers, Ed. (Springer, 2009), pp. 1596–1615.
21. J. D. Stevenson, J. Schmalian, P. G. Wolynes, *Nat. Phys.* **2**, 268–274 (2006).
22. G. Biroli, J.-P. Bouchaud, in *Structural Glasses and Supercooled Liquids: Theory, Experiment, and Applications*, P. G. Wolynes, V. Lubchenko, Eds. (Wiley, 2012), pp. 31–114.
23. F. Ladieu, C. Brun, D. L'Hôte, *Phys. Rev. B* **85**, 184207 (2012).
24. R. Richert, S. Weinstein, *Phys. Rev. Lett.* **97**, 095703 (2006).
25. W. T. Coffey, B. V. Paranjape, *Proc. R. Irish Acad. A* **78**, 17–25 (1978).
26. T. Blochowicz, C. Tschirwitz, S. Benkhof, E. A. Rössler, *J. Chem. Phys.* **118**, 7544–7555 (2003).
27. P. G. Debenedetti, F. H. Stillinger, *Nature* **410**, 259–267 (2001).
28. R. Böhmer, K. L. Ngai, C. A. Angell, D. J. Plazek, *J. Chem. Phys.* **99**, 4201–4209 (1993).

ACKNOWLEDGMENTS

We thank C. Alba-Simionesco, A. Coniglio, P.-M. Déjardin, G. Tarjus, and M. Tarzia for interesting discussions. The work in Saclay was supported by ERC grant NPRGLASS, by the Labex RTRA grant Aricover, and by the Institut des Systèmes Complexes ISC-PIF. The work in Augsburg was supported by the Deutsche Forschungsgemeinschaft via Research Unit FOR1394. S.A., R.T., C.W.-G., and F.L. developed the experimental setup at Saclay; Th.B. and P.L. developed the experimental setup at Augsburg; Th.B. and M.M. performed the measurements and analysis of the Augsburg data; A.L. and P.L. conceived and supervised the project in Augsburg; G.B. and J.-P.B. derived the theoretical scaling analysis; and all authors were involved in the interpretation of the results and creation of this manuscript. Data are available as supplementary material. All authors have no competing financial interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1308/suppl/DC1
Materials and Methods
Figs. S1 to S3
References (*29*–*42*)
Data files

25 January 2016; accepted 9 May 2016
10.1126/science.aaf3182

CARBON SEQUESTRATION

Rapid carbon mineralization for permanent disposal of anthropogenic carbon dioxide emissions

Juerg M. Matter,[1,2]* Martin Stute,[2] Sandra Ó. Snæbjörnsdottir,[3] Eric H. Oelkers,[3,4,5] Sigurdur R. Gislason,[3] Edda S. Aradottir,[6] Bergur Sigfusson,[6,7] Ingvi Gunnarsson,[6] Holmfridur Sigurdardottir,[6] Einar Gunnlaugsson,[6] Gudni Axelsson,[8] Helgi A. Alfredsson,[3] Domenik Wolff-Boenisch,[3,9] Kiflom Mesfin,[3] Diana Fernandez de la Reguera Taya,[2] Jennifer Hall,[2] Knud Dideriksen,[10] Wallace S. Broecker[2]

Carbon capture and storage (CCS) provides a solution toward decarbonization of the global economy. The success of this solution depends on the ability to safely and permanently store CO_2. This study demonstrates for the first time the permanent disposal of CO_2 as environmentally benign carbonate minerals in basaltic rocks. We find that over 95% of the CO_2 injected into the CarbFix site in Iceland was mineralized to carbonate minerals in less than 2 years. This result contrasts with the common view that the immobilization of CO_2 as carbonate minerals within geologic reservoirs takes several hundreds to thousands of years. Our results, therefore, demonstrate that the safe long-term storage of anthropogenic CO_2 emissions through mineralization can be far faster than previously postulated.

The success of geologic CO_2 storage depends on its long-term security and public acceptance, in addition to regulatory, policy, and economical factors (*1*). CO_2 and brine leakage through a confining system above the storage reservoir or through abandoned wells is considered one of the major challenges associated with geologic CO_2 storage (*2*–*4*). Leakage rates into the atmosphere of ≤0.1% are required to ensure effective climate change mitigation (*5*, *6*). To avoid CO_2 leakage, caprock integrity needs to be evaluated and monitored (*7*). Leakage risk is further enhanced by induced seismicity, which may open fluid flow pathways in the caprock (*8*). Mineral carbonatization (i.e., the conversion of CO_2 to carbonate minerals) via CO_2-fluid-rock reactions in the reservoir minimizes the risk of leakage and thus facilitates long-term and safe carbon storage and public acceptance (*9*). The potential for carbonatization is, however, limited in conventional CO_2 storage reservoirs such as deep saline aquifers and depleted oil and gas reservoirs in sedimentary basins due to the lack of calcium-, magnesium-, and iron-rich silicate minerals required to form carbonate minerals (*10*, *11*). An alternative is to inject CO_2 into basaltic rocks, which contain up to 25% by weight of calcium, magnesium, and iron. Basaltic rocks are highly reactive and are one of the most common rock types on Earth, covering ~10% of continental surface area and most of the ocean floor (*12*, *13*).

The CarbFix pilot project in Iceland was designed to promote and verify in situ CO_2 mineralization in basaltic rocks for the permanent disposal of anthropogenic CO_2 emissions (*14*). Two injection tests were performed at the CarbFix injection site near the Hellisheidi geothermal power plant. Phase I: 175 tons of pure CO_2 from January to March 2012, and phase II: 73 tons of a CO_2-H_2S gas mixture in June to August 2012, of which 55 tons were CO_2. H_2S is not only a major constituent of geothermal gases but also of CO_2-rich sour gas. Because the cost of carbon capture and storage (CCS) is dominated by the cost of capture and gas separation, the overall cost could be lowered substantially by injecting gas mixtures rather than pure CO_2 (*9*). Hence, the purpose of the mixed CO_2/H_2S injection was to assess the feasibility of injecting impurities in the CO_2 stream.

The CarbFix injection site is situated about 25 km east of Reykjavik and is equipped with a 2000-m-deep injection well (HN02) and eight monitoring wells ranging in depth from 150 to 1300 m (Fig. 1). The target CO_2 storage formation is at between 400 and 800 m depth and consists of basaltic lavas and hyaloclastites with lateral and vertical intrinsic permeabilities of 300 and 1700 × 10^{-15} m^2, respectively (*15*, *16*). It is overlain by low-permeability hyaloclastites. The formation water temperature and pH in the injection interval range from 20° to 33°C and from 8.4 to 9.4, and it is oxygen-depleted (*15*). Due to the shallow depth of the target storage reservoir and the risk of CO_2 gas leakage through fractures, a novel CO_2 injection system was designed and used, which dissolves the gases into down-flowing water in the well during its injection (*17*). To avoid potential degassing, the CO_2 concentration in the injected fluids was kept below its solubility at reservoir conditions (*17*). Once dissolved in water, CO_2 is no longer buoyant (*17*), and it immediately starts to react with the Ca-Mg-Fe–rich reservoir rocks.

Because dissolved or mineralized CO_2 cannot be detected by conventional monitoring methods such as seismic imaging, the fate of the injected CO_2 was monitored with a suite of chemical and isotopic tracers. The injected CO_2 was spiked with carbon-14 (^{14}C) to monitor its transport and reactivity (*18*). For the pure CO_2 and the CO_2/H_2S injections, the ^{14}C concentrations of the injected fluids were 40.0 Bq/liter ($^{14}C/^{12}C$: 2.16×10^{-11}) and 6 Bq/liter ($^{14}C/^{12}C$: 6.5×10^{-12}), respectively. By comparison, the ^{14}C concentration in the reservoir before the injections was 0.0006 Bq/liter ($^{14}C/^{12}C$: 1.68×10^{-13}). This novel carbon tracking method was previously proposed for geologic CO_2 storage monitoring, but its feasibility has not been tested previously (*19*, *20*). Because $^{14}CO_2$ chemically and physically behaves identically to $^{12}CO_2$ and is only minimally affected by isotope fraction during phase transitions (*21*), it provides the means to accurately inventory the fate of the injected carbon.

In addition to ^{14}C, we continuously co-injected nonreactive but volatile sulfur hexafluoride (SF_6) and trifluoromethyl sulfur pentafluoride (SF_5CF_3) tracers to assess plume migration in the reservoir. The SF_6 was used during phase I and SF_5CF_3 during phase II. The SF_6 and SF_5CF_3 concentrations in the injected fluids were 2.33×10^{-8} cc at standard temperature and pressure (ccSTP)/cc and 2.24×10^{-8} ccSTP/cc, respectively.

The CO_2 and CO_2/H_2S mixtures, together with the tracers, were injected into the target storage formation fully dissolved in water pumped from a nearby well. Typical injection rates during phase I injection were 70 g/s for CO_2 and 1800 g/s for H_2O, respectively (*17*). Injection rates during phase II varied between 10 and 50 g/s for CO_2 and 417 and 2082 g/s for H_2O. The dissolved inorganic carbon (DIC) concentration and pH of the injectates were 0.82 mol/liter and 3.85 (at 20°C) for phase I and 0.43 mol/liter and 4.03 for phase II. Fluid samples for SF_6, SF_5CF_3, ^{14}C, DIC, and pH analyses were collected without degassing using a specially designed downhole sampler from the injection well HN02 (*22*) or with a submersible pump from the first monitoring well, HN04, located about 70 m downstream from HN02 at 400 m depth below the surface before, during, and after injection (tables S1 to S3).

The arrival of the injectate from phase I at monitoring well HN04 was confirmed by an increase in SF_6 concentration, and a sharp decrease in pH and DIC concentration (Fig. 2, A and B, and table S3). Based on the SF_6 data, the initial breakthrough in HN04 occurred 56 days after injection. Subsequently, the SF_6 concentration slightly decreased before a further increase in concentration occurred, with peak concentration 406 days after initiation of the injection. SF_5CF_3 behaves similarly (Fig. 2A); its initial arrival was detected 58 days after initiation of the phase II injection,

[1]Department of Ocean and Earth Science, University of Southampton, Southampton, UK. [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA. [3]Institute of Earth Sciences, University of Iceland, Iceland. [4]CNRS/UMR 5563, Université Paul Sabatier, Toulouse, France. [5]Department of Earth Science, University College London, UK. [6]Reykjavik Energy, Reykjavik, Iceland. [7]European Commission, Joint Research Center, Institute for Energy and Transport, Petten, Netherlands. [8]Iceland GeoSurvey, Reykjavik, Iceland. [9]Department of Applied Geology, Curtin University, Perth, Western Australia. [10]Nano-Science Center, Department of Chemistry, University of Copenhagen, Copenhagen, Denmark.
***Corresponding author. Email: j.matter@southampton.ac.uk**

followed by decreasing concentrations until 350 days after the injection started. Subsequently, the SF_5CF_3 concentration increased, consistent with the SF_6 tracer breakthrough curve. The double peaks in these tracer breakthrough curves are also in agreement with results from previous tracer tests showing that the storage formation consists of relatively homogenous porous media intersected by a low-volume and fast flow path that channels about 3% of the tracer flow between HN02 and HN04 (*23*).

The time series of DIC, pH, and ^{14}C in HN04 are initially coincident with the SF_6 record, showing peak concentrations in ^{14}C and DIC and a decrease in pH around 56 days after injection (Figs. 2B and 3). The small drop in pH and increase in DIC around 200 days after injection is caused by the phase II injection, as confirmed by the SF_5CF_3 time series (Fig. 2A). The similar initial pattern in the tracer breakthrough curves and the DIC concentration suggests identical transport behavior of carbon and tracers in the reservoir. However, ^{14}C and DIC concentrations subsequently decreased and stayed more or less constant for the remaining monitoring period, with the exception of a small increase in concentration induced by the phase II injection (Figs. 2B and 3, A and B).

The fate of the injected CO_2 was quantified using mass balance calculations (*18*). The resulting calculated DIC and ^{14}C concentrations are much higher than those measured in the collected water samples, suggesting a loss of DIC and ^{14}C along the subsurface flow path toward the monitoring well (Fig. 3, A and B). The most plausible mechanism for this difference is carbonate precipitation. The differences between calculated and measured DIC and ^{14}C indicate that >95% of the injected CO_2 was mineralized through water-CO_2-basalt reactions between the injection (HN02) and monitoring (HN04) wells within 2 years (Fig. 3, A and B). The initial peak concentrations in DIC and ^{14}C detected around 56 days after injection suggest that travel time along the low-volume fast-flowing flow path was too short for significant CO_2 mineralization to occur. Most of the injected CO_2 was probably mineralized within the porous matrix of the basalt that allows for longer fluid residence times and thus extended reaction time. This conclusion is confirmed by (i) calculated fluid saturation states showing that the collected monitoring fluids are at saturation or supersaturation with respect to calcite at all times except during the initial low-volume flow path contribution; (ii) x-ray diffraction and scanning electron microscopy with energy-dispersive x-ray spectroscopy analysis of secondary mineral precipitates collected from the submersible pump in monitoring well HN04 after it was hauled to the surface, showing these precipitates to be calcite (*18*) (figs. S1 to S3); and (iii) the similarity in the ^{14}C concentration of the injected CO_2 and the precipitated collected calcite (7.48 ± 0.8 and 7.82 ± 0.05 fraction modern).

Although monitoring continues, the time scale of the tracer and DIC data discussed is limited to 550 days, because most of the injected CO_2 was mineralized by this time (Figs. 2 and 3). This 550-day

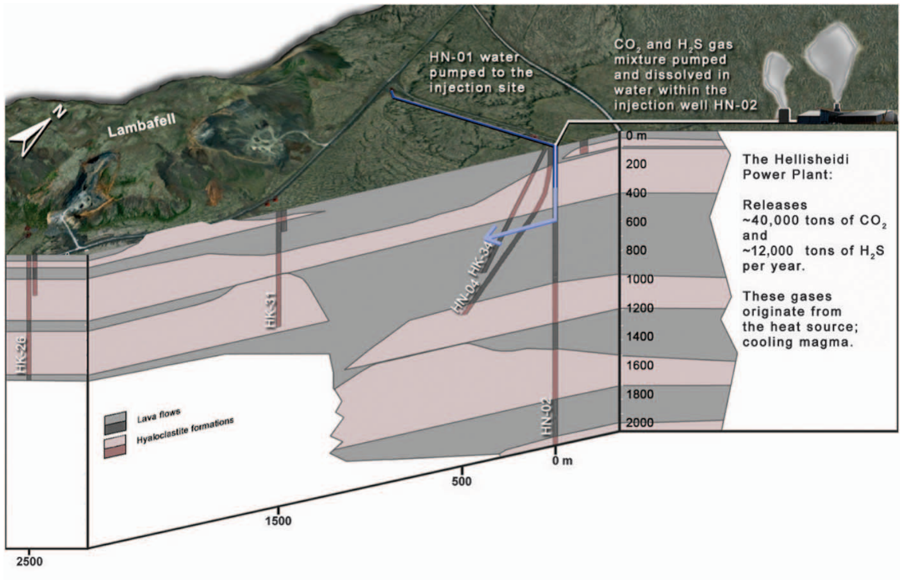


Fig. 1. Geological cross-section of the CarbFix injection site. CO_2 and H_2S are injected fully dissolved in water in injection well HN02 at a depth between 400 and 540 m. For this study, fluid samples were collected in the injection well HN02 and the monitoring well HN04 [modified from (*15*)].

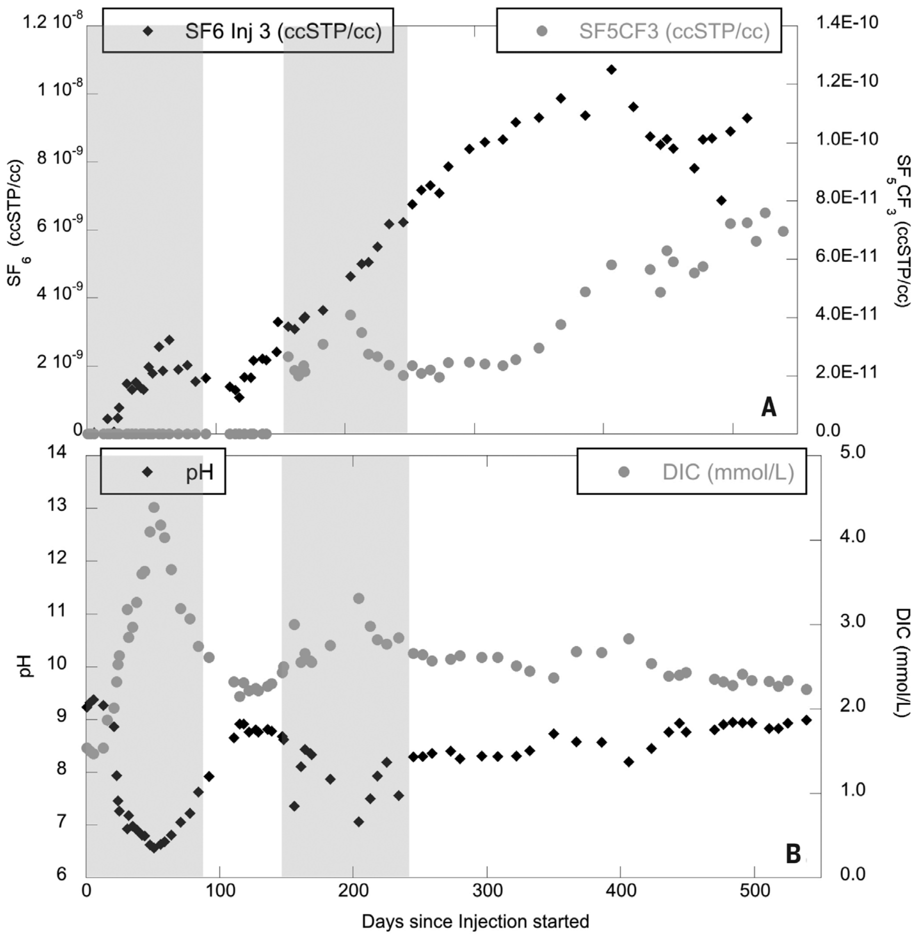


Fig. 2. Change of tracer concentrations, DIC, and pH in the target CO_2 storage formation fluid. Time series of (**A**) SF_6 and SF_5CF_3 tracer concentrations (ccSTP/cc) and (**B**) pH and DIC in monitoring well HN04 for the pure CO_2 and the CO_2 and H_2S injections. The shaded area indicates the phase I and II injection period.

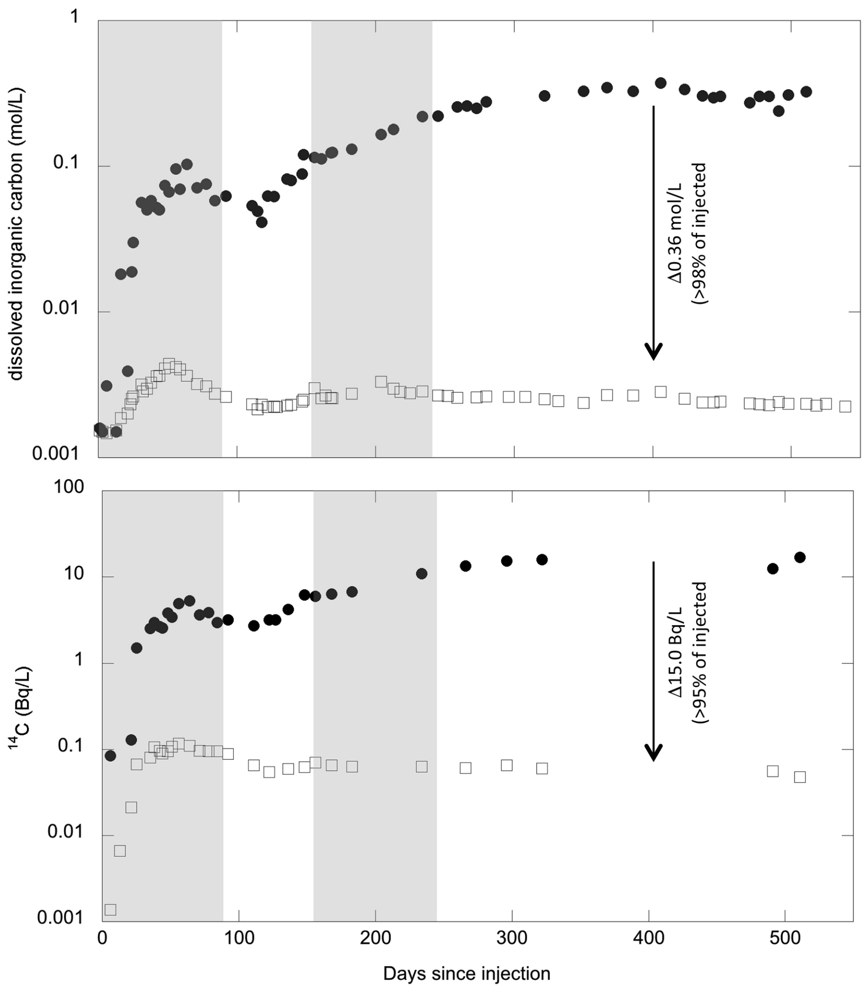


Fig. 3. Comparison of calculated and measured DIC and ^{14}C concentrations in the target CO_2 storage formation fluid. (**A**) Time series of expected (solid circles) versus measured (open squares) DIC (mol/liter) in monitoring well HN04, indicating >98% conversion of injected CO_2 to carbonate minerals, and (**B**) time series of expected (solid circles) versus measured (open squares) $^{14}C_{DIC}$ (Bq/liter) in monitoring well HN04, showing >95% of injected CO_2 to be converted to carbonate minerals. The shaded area indicates the phase I and II injection periods.

limit also coincides with the breakdown of the submersible pump in HN04 monitoring well, which resulted in a 3-month gap in the subsequent monitoring data. The pump was clogged and coated with calcite (*18*).

The fast conversion rate of dissolved CO_2 to calcite minerals in the CarbFix storage reservoir is most likely the result of several key processes: (i) the novel CO_2 injection system that injected water-dissolved CO_2 into the subsurface; (ii) the relatively rapid dissolution rate of basalt, releasing Ca, Mg, and Fe ions required for the CO_2 mineralization; (iii) the mixing of injected water with alkaline formation waters; and (iv) The dissolution of preexisting secondary carbonates at the onset of the CO_2 injection, which may have contributed to the neutralization of the injected CO_2-rich water via the reaction $CaCO_3 + CO_2 + H_2O = Ca^{2+} + 2\ HCO_3^-$.

The dissolution of preexisting calcite is supported by the $^{14}C/^{12}C$ ratio of the collected fluid samples, which suggest a 50% dilution of the carbon in the fluid, most likely via calcite dissolution just after it arrives in the basaltic reservoir. Nevertheless, the mass balance calculations clearly demonstrate that these preexisting carbonates re-precipitated during the mineralization of the injected CO_2.

The results of this study demonstrate that nearly complete in situ CO_2 mineralization in basaltic rocks can occur in less than 2 years. Once stored within carbonate minerals, the leakage risk is eliminated and any monitoring program of the storage site can be significantly reduced, thus enhancing storage security and potentially public acceptance. Natural aqueous fluids in basalts and those at the CarbFix site tend to be at or close to equilibrium with respect to calcite, limiting its redissolution (*16*). The scaling up of this basaltic carbon storage method requires substantial quantities of water and porous basaltic rocks (*9*). Both are widely available on the continental margins, such as off the coast of the Pacific Northwest of the United States (*12*).

REFERENCES AND NOTES

1. B. Metz, O. Davidson, H. de Coninck, M. Loos, L. A. Meyer, Eds., *IPCC Special Report on Carbon Dioxide Capture and Storage* (Cambridge Univ. Press, New York, 2005).
2. A. Esposito, S. Benson, in Proceedings of the Society of Petroleum Engineers 2010 Western North American Regional Meeting, Anaheim, CA, 27 to 29 May 2010 (Society of Petroleum Engineers, 2010), SPE-133604.
3. B. Ellis *et al*., *Sci. Technol.* **1**, 248–260 (2011).
4. J. M. Bielicki, M. F. Pollak, J. P. Fitts, C. A. Peters, E. J. Wilson, *Int. J. Greenhouse Gas Control* **20**, 272–284 (2014).
5. P. M. Haugan, F. Joos, *Geophys. Res. Lett.* **31**, L18202 (2004).
6. B. van der Zwaan, L. Smekens, *Environ. Model. Assess.* **14**, 135–148 (2009).
7. H. S. Eggleston *et al*., Eds., *IPCC Guidelines for National Greenhouse Gas Inventories – A Primer* (National Greenhouse Gas Inventories Programme, Institute for Global Environmental Strategies, Japan, 2008).
8. M. D. Zoback, S. M. Gorelick, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10164–10168 (2012).
9. S. R. Gislason, E. H. Oelkers, *Science* **344**, 373–374 (2014).
10. J. M. Matter, P. B. Kelemen, *Nat. Geosci.* **2**, 837–841 (2009).
11. S. M. V. Gilfillan *et al*., *Nature* **458**, 614–618 (2009).
12. D. S. Goldberg, T. Takahashi, A. L. Slagle, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9920–9925 (2008).
13. B. P. McGrail *et al*., *J. Geophys. Res.* **111**, B12201 (2006).
14. S. R. Gislason *et al*., *Int. J. Greenhouse Gas Control* **4**, 537–545 (2010).
15. H. A. Alfredsson *et al*., *Int. J. Greenhouse Gas Control* **12**, 399–418 (2013).
16. E. S. Aradóttir, E. L. Sonnenthal, G. Björnsson, H. Jónsson, *Int. J. Greenhouse Gas Control* **9**, 24–40 (2012).
17. B. Sigfusson *et al*., *Int. J. Greenhouse Gas Control* **37**, 213–219 (2015).
18. Materials and methods are available as supplementary materials on *Science* Online.
19. P. P. Bachelor *et al*., *J. Radioanal. Nucl. Chem.* **277**, 85–89 (2008).
20. K. S. Lackner, S. Brennan, *Clim. Change* **96**, 357–378 (2009).
21. I. D. Clark, P. Fritz, *Environmental Isotopes in Hydrogeology* (Lewis Publishers, Boca Raton, FL, 1997).
22. H. A. Alfredsson, K. Mesfin, D. Wolff-Boenisch, *Greenhouse Gas Sci. Technol.* **6**, 167–177 (2016).
23. M. Rezvani Khalilabad, G. Axelsson, S. Gislason, *Mineral. Mag.* **72**, 121–125 (2008).

ACKNOWLEDGMENTS

We acknowledge funding from Reykjavik Energy; the U.S. Department of Energy under award number DE-FE0004847 to J.M.M and M.S.; the European Commission through the projects CarbFix (EC coordinated action 283148), Min-GRO (MC-RTN-35488), Delta-Min (PITN-GA-2008-215360), and CO_2-REACT (EC Project 317235) to S.R.G., E.H.O., and Reykjavik Energy; Nordic fund 11029-NORDICCS; and the Icelandic GEORG Geothermal Research fund (09-02-001) to S.R.G. and Reykjavik Energy. We thank T. Kristinsson and E. Ö. Þrastarsonfor for helping with sample collection in the field. All data used in this study are included in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1312/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S3
Tables S1 and S2
References (*24–34*)

10.1126/science.aad8132

SLEEP AND MEMORY

Top-down cortical input during NREM sleep consolidates perceptual memory

D. Miyamoto,[1,2,3,4] D. Hirai,[1] C. C. A. Fung,[5] A. Inutsuka,[2] M. Odagawa,[1] T. Suzuki,[1] R. Boehringer,[6] C. Adaikkan,[6] C. Matsubara,[1] N. Matsuki,[3] T. Fukai,[5] T. J. McHugh,[6] A. Yamanaka,[2] M. Murayama[1]*

During tactile perception, long-range intracortical top-down axonal projections are essential for processing sensory information. Whether these projections regulate sleep-dependent long-term memory consolidation is unknown. We altered top-down inputs from higher-order cortex to sensory cortex during sleep and examined the consolidation of memories acquired earlier during awake texture perception. Mice learned novel textures and consolidated them during sleep. Within the first hour of non–rapid eye movement (NREM) sleep, optogenetic inhibition of top-down projecting axons from secondary motor cortex (M2) to primary somatosensory cortex (S1) impaired sleep-dependent reactivation of S1 neurons and memory consolidation. In NREM sleep and sleep-deprivation states, closed-loop asynchronous or synchronous M2-S1 coactivation, respectively, reduced or prolonged memory retention. Top-down cortical information flow in NREM sleep is thus required for perceptual memory consolidation.

Non–rapid eye movement (NREM) sleep is essential for memory consolidation of an animal's awake motor (*1*) and sensory (*2*) learning experiences. During NREM, synchronous 0.5 to 4 Hz oscillations (slow-wave activity, SWA) sweep across cortical areas (*3*–*5*) suggesting that interregional transfer of internal information in NREM has a role in memory consolidation (*6*–*8*). We recently identified a reverberating long-range top-down intracortical circuit that underlies somatosensory perception in the mouse hindpaw (*9*), consisting of sensory input from the primary somatosensory cortex (S1) to the secondary motor cortex (M2) and a reciprocal top-down feedback projection from M2 to S1. Optogenetic inhibition of the top-down projection impaired accurate tactile perception. However, whether similar top-down cortical inputs have a critical role in memory consolidation during sleep remains unexamined.

To assess memory consolidation, we developed a floor-texture recognition (FTR) task (Fig. 1A) based on natural novelty preference in mice (*10*). During the sampling period, mice explored objects on the left and right sides of an arena containing smooth floors with no behavioral preference (Fig. 1B). However, during the testing period, mice preferentially explored the object on the novel texture (Fig. 1B). We defined the strength of memory for the familiar texture by the relative amount of time spent exploring an object on the novel texture and for a second texture combination (fig. S2).

The memory retention period lasted for 2 days (fig. S3). Mice do not have an innate preference for groovy or smooth textures (fig. S4). No change in performance was observed with whisker-trimmed mice in a dark room (fig. S5). Optogenetic silencing of S1 sensory cortex hindpaw area impaired task performance (fig. S6). Similar to other perceptual recognition tasks (*11*), behavioral performance declined with sleep deprivation (SD) after the sampling period (fig. S7). In the light phase, SD during the first hour of the resting period, after the sampling period, produced a decline in performance that was not observed with SD in the resting period 6 to 7 hours later (fig. S7). Even in testing periods starting immediately after SD (i.e., 1-hour interval with SD), mice showed impaired consolidation (fig. S7). Without SD, mice showed impaired consolidation at the start of the dark phase (fig. S7), when mice are in an active period and in a shorter sleep period (*12*). These results indicate the sleep dependence of the FTR task. Recognition tasks may involve synaptic plasticity (*13*), and *N*-methyl-D-aspartate (NMDA) receptor blockers degraded performance (fig. S8). To test hippocampal dependence, we injected adeno-associated virus (AAV)–FLEX–tetanus toxin (*14*) into the CA1 region of calcium/calmodulin-dependent protein kinase IIα–Cre transgenic mice (*15*) and immunohistochemically confirmed the blockade of synaptic transmission (*16*) at the subiculum, the primary site of termination of CA1 axons (fig. S9). Performance in the FTR task did not decline, which indicated that it is likely independent of the hippocampus (fig. S9).

We examined whether M2 input to S1 affected ongoing perception and consolidation by optogenetically inactivating neural activity during the sampling period, resting period, or testing period in the FTR task (fig. S10). Inactivation of M2 fibers in S1 during the sampling or testing periods decreased task performance (fig. S10). S1 firing activity in the awake state was significantly reduced with top-down inactivation (86.8 ± 2.8%, n = 24 units, P < 0.001, light on versus off periods, one-sample *t*-test). Inactivated S1 fibers in M2 resulted in similar data (fig. S10).

We performed optogenetic inactivation of M2 fibers in S1 (Fig. 2A) during the resting period in the first hour after the sampling period that included at least three brain states: 42.9 ± 6.3% (n = 7 mice) in an awake state (normalized to 1 hour) with small high-frequency electroencephalographic (EEG) activity and large electromyographic (EMG) activity; 54.5 ± 5.9% (n = 7) in NREM sleep with SWA of 0.5 to 4 Hz and a silent EMG; and 2.6 ± 0.96% (n = 7) in rapid eye movement (REM) sleep with EEG activity similar to the awake state and a silent EMG (Fig. 2B). We inactivated M2 fibers or S1 fibers during the resting period in the first hour or 6 to 7 hours after the sampling period (Fig. 2C) with a closed-loop online photostimulation system (>90% accuracy) (see fig. S11). A quiet wakefulness (QW) state was negligible in the first hour (0.4 ± 0.1% in NREM illumination, 3.4 ± 1.2% in awake illumination) (fig. S11). Photostimulation in the awake or NREM sleep states did not alter the total duration of the three brain states (fig. S12 and S13). Optogenetic inactivation of M2 fibers during the resting-awake state did not affect task performance in the testing period (Fig. 2D) but significantly decreased task performance during the resting period–NREM sleep immediately after the sampling period (Fig. 2D), whereas inactivation during resting-NREM sleep 6 to 7 hours after the sampling period did not (Fig. 2D). Similarly, inactivation of S1 fibers during resting-NREM sleep immediately after (0 to 1 hour) the sampling period did not alter performance (Fig. 2E). In contrast, a visual-based task was not affected by optogenetic inactivation of hindpaw M2 to S1 inputs in the NREM periods (fig. S14). These results indicated that tactile memory consolidation requires M2 input to S1 during NREM sleep shortly after the sampling period.

We asked why M2 to S1 top-down input, and not vice versa, regulates memory consolidation specifically during resting-NREM sleep (Fig. 2D). We hypothesized that the optogenetics disrupts the causal directional regulation by M2 to S1 activity and/or suppresses the prominent emergence of reactivated neurons in S1, as SWA during NREM sleep propagates in an anteroposterior direction (*3*, *4*, *17*). NREM sleep accompanies the reactivation of neurons related to an animal's sensory experience before sleep, which is thought to be crucial for memory consolidation (*18*). We performed a Granger causality analysis (*19*, *20*), which can predict directed functional (causal) connections between cortical areas. We recorded and analyzed local field potentials (LFPs) from both M2 and S1 with and without optogenetic inactivation of the M2 to S1 projection (Fig. 3, A to C) (note: hereafter, we focused on the resting

[1]Laboratory for Behavioral Neurophysiology, RIKEN Brain Science Institute, Wako, Saitama, Japan. [2]Department of Neuroscience II, Research Institute of Environmental Medicine, Nagoya University, Nagoya, Japan. [3]Laboratory of Chemical Pharmacology, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan. [4]Japan Society for the Promotion of Science Research Fellow, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo, 102-0083, Japan [5]Laboratory for Neural Circuit Theory, RIKEN Brain Science Institute, Wako, Saitama, Japan. [6]Laboratory for Circuit and Behavioral Physiology, RIKEN Brain Science Institute, Wako, Saitama, Japan.
***Corresponding author. Email: masa_murayama@brain.riken.jp**

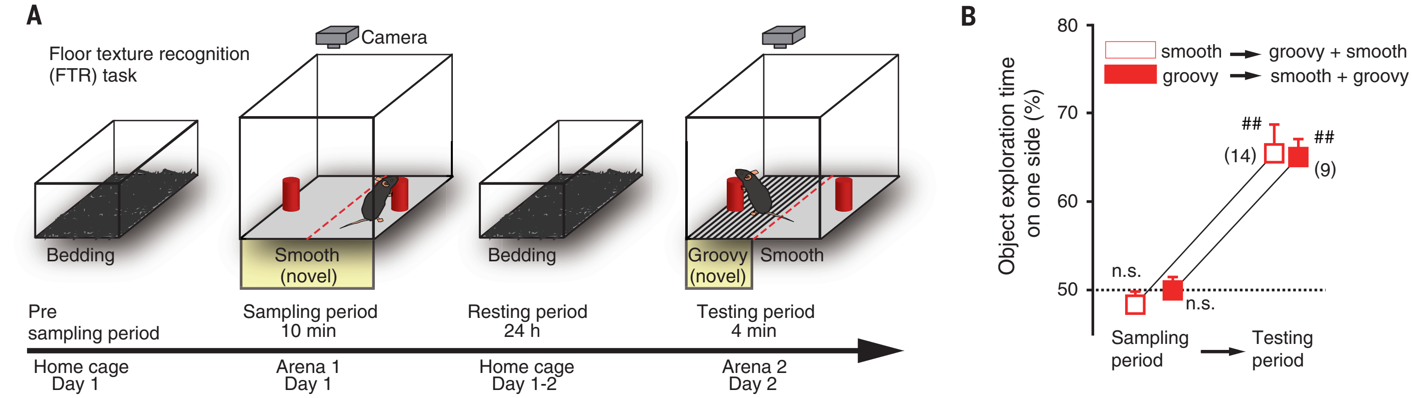
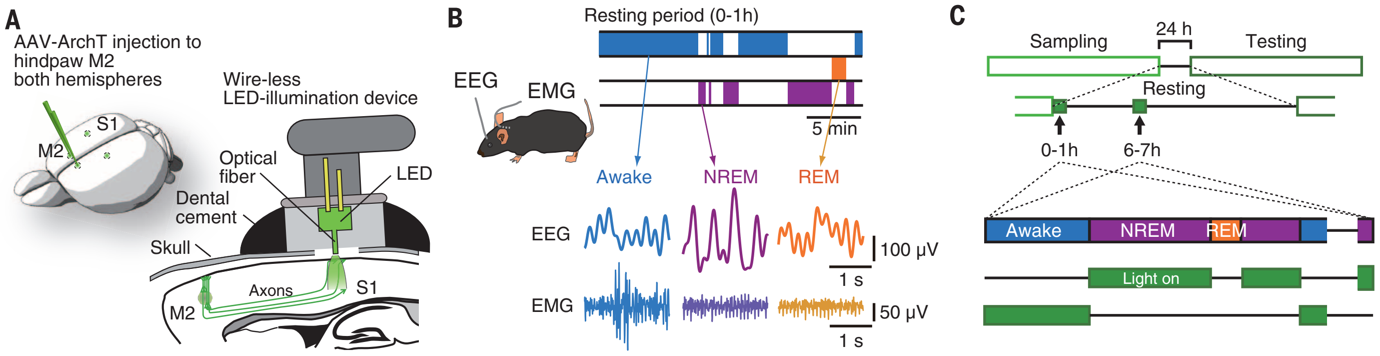


Fig. 1. Floor texture recognition task. (**A**) Behavioral paradigm. (**B**) Object exploration time on a texture in the sampling period and on a novel texture in the testing period. Data are means ± SEM; statistical significance from 50% chance level (##*P* < 0.01) was assessed by one-sample *t* test. n.s., Not significant.



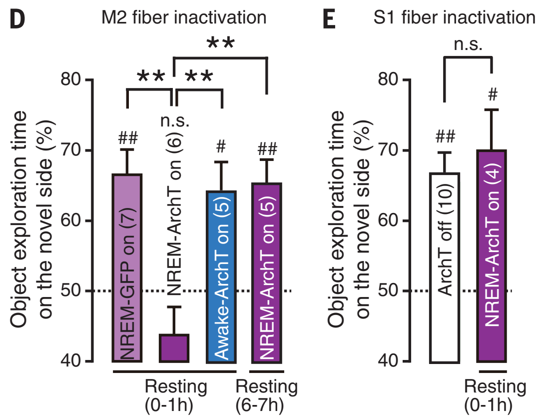


Fig. 2. Optogenetic inactivation of M2 axons impairs memory consolidation. (**A**) Diagram of the miniature wireless light-emitting diode device that was attached to S1 (or M2) in both hemispheres. AAV-ArchT or AAV-GFP (green fluorescent protein) was injected (inset) into M2 (or S1) in both hemispheres. (**B**) Examples of EEG and EMG recordings during the resting period. Brain states were identified with EEG recordings (see Methods). (**C**) Diagram of sleep state–specific optogenetics. (**D**) Summary for the task when M2 fibers were inactivated at S1 during the three periods. (**E**) Summary for the task when S1 fibers were inactivated at M2 during resting-NREM sleep (0 to 1 hour after sampling period). The cumulative illumination time was 30 min in each state. (D) and (E), *n* in parentheses. Statistical significance among more than two groups (***P* < 0.01) was assessed by one-way analysis of variance (ANOVA) with Tukey's post hoc test, statistical significance between two groups was assessed by Student's *t* test, statistical significance from 50% chance level (#*P* < 0.05, ##*P* < 0.01) was assessed by one-sample *t* test.

period immediately after the sampling period). Brain state–dependent photostimulation was performed by visual observation of online EEG and EMG data (fig. S11). Coherence analysis indicated higher synchronized M2–S1 activity in the delta range (0.5 to 4Hz) in S1 and M2 LFPs during NREM sleep compared with the awake state (Fig. 3D). A power spectrum in the delta range during NREM sleep was not altered by optogenetic inactivation (Fig. 3E and fig. S15). Without optogenetics, we noted a significant increase in causality in both directions (M2 to S1 and vice versa) during the sampling period and less causality during the resting-awake state (fig. S16), which suggested that ongoing perception requires both pathways (fig. S10) (*9*). The lower causality from M2 to S1 suggests that memory consolidation does not require M2 inputs during the resting-awake state (Fig. 2, D). Causality from M2 to S1 during presampling-NREM sleep and resting-NREM sleep were significantly higher than the resting-awake state, which indicated brain state dependence (fig. S17). With optogenetic inactivation, we observed a decrease of causality only in the M2→S1 direction during resting-NREM sleep (Fig. 3F).

We tested whether reactivated neurons were suppressed by optogenetic inactivation with tetrode recordings from M2 and S1 (Fig. 3G). Single-unit activity was recorded during the sampling period, and NREM sleep before (pre-) and after (post-) the sampling (fig. S18). Reactivated neurons were defined on the basis of total firing activity normalized to presampling NREM sleep (*1*). S1 and M2 neurons that were active in the sampling period were also active during postsampling NREM sleep. Conversely, neurons that were less active in the sampling period were also less active (Fig. 3, H and I). Optogenetic inactivation of M2 fibers did not suppress this linear correlation in M2 (Fig. 3H), but decreased it in S1 (Fig. 3I). Optogenetic inactivation of M2 axons suppressed M2→S1 causality (Fig. 3F), reactivated S1 neurons (Fig. 3I), and task performance (Fig. 2D).

We tested whether neuronal reactivation is sufficient for memory consolidation by optogenetically activating both areas in a synchronous or antisynchronous manner (Fig. 4A). We used Thy1-channelrhodopsin-2 (ChR2) transgenic mice in which mostly L5 cortical neurons express ChR2 (*21*). We confirmed that light stimulation to M2 and S1 reliably evokes firing during NREM sleep

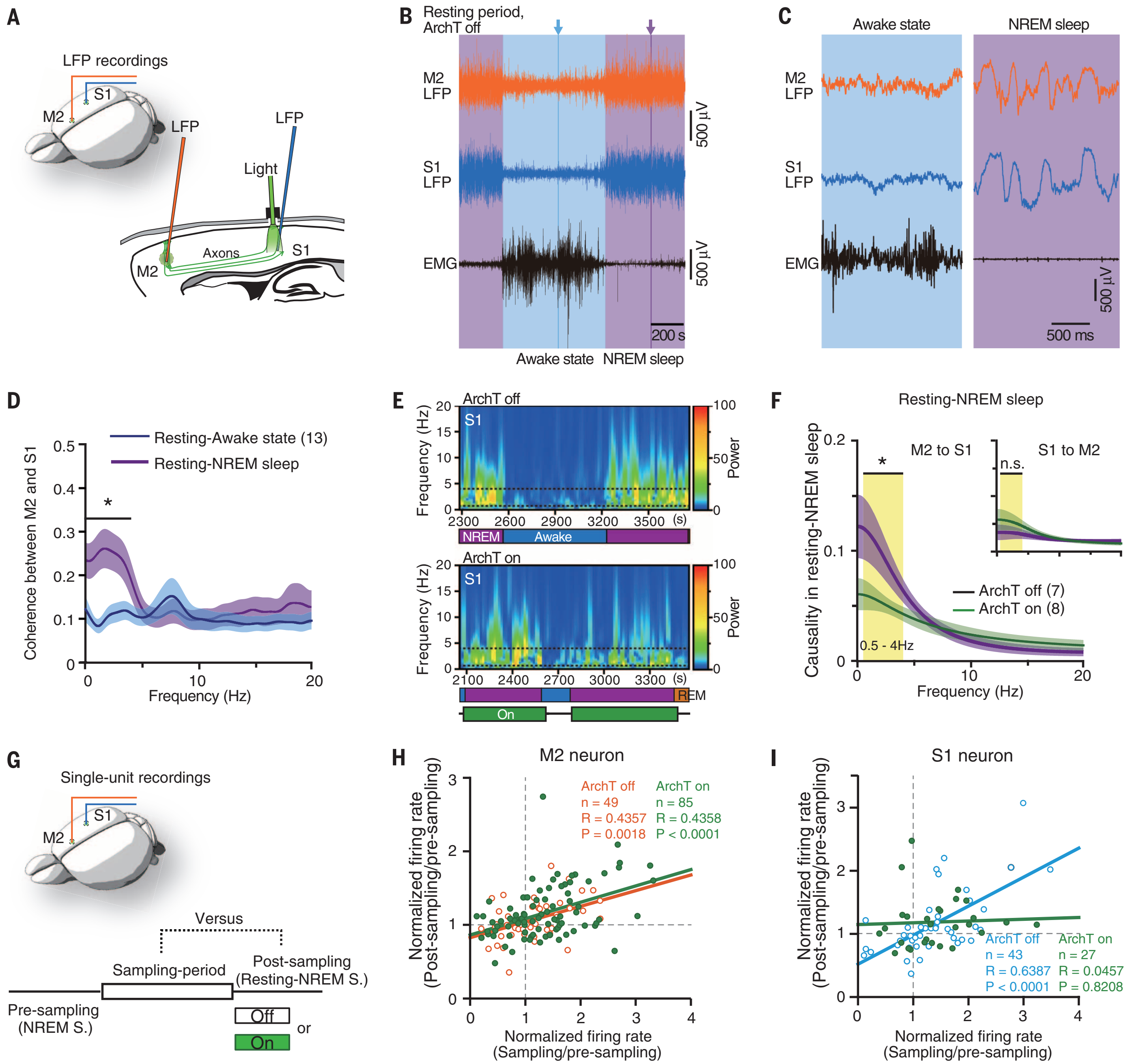
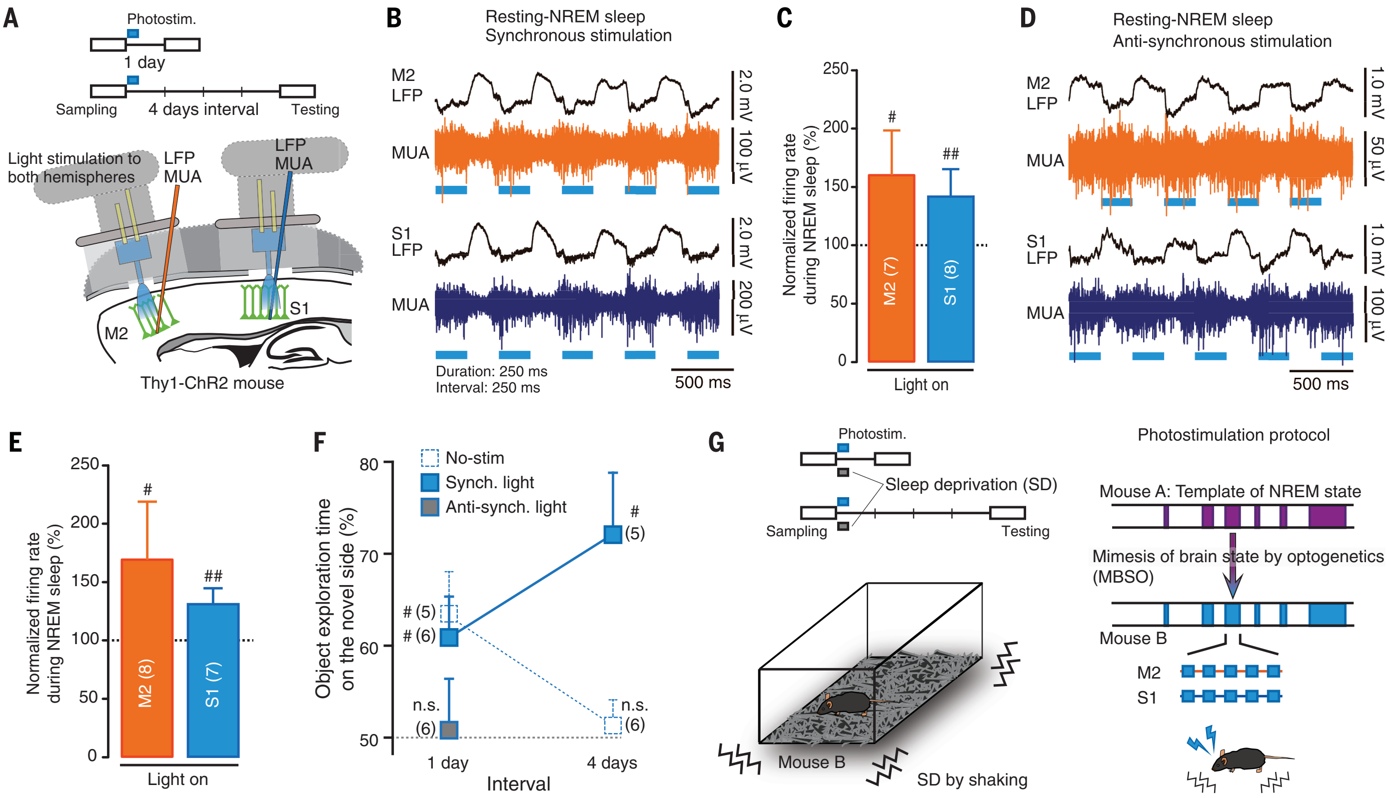


Fig. 3. Optogenetic inactivation of M2 fibers during NREM sleep suppressed M2-S1 causality and reactivated S1 neurons. (**A**) Recording of local field potentials (LFPs) from M2 and S1 with fiberoptic illumination. (**B** and **C**) Representative data of LFPs and EMG (**C**), expanded trace at arrows (**B**) during resting-awake and NREM states without optogenetics. (**D**) Coherence between S1 and M2 during resting-awake and NREM states. (**E**) Representative frequency spectrum of LFPs with and without optogenetic inhibition of M2 fibers at S1. NREM sleep identified by post hoc analysis and online visual observation–based photostimulation (green) are shown at the bottom (see fig. S11 for accuracy). (**F**) Granger causality during the resting-NREM periods with and without optogenetics. (**G**) Diagram of tetrode recordings and fiberoptic illumination. Single-unit activities were collected during presampling (resting-NREM) for normalization of the sampling and postsampling (resting-NREM) periods. (**H** and **I**) Normalized firing rate from M2 and S1 individual units (each circle) during the sampling and postsampling (resting-NREM) periods with (green) and without (black) optogenetic M2 fiber inactivation at S1. Statistical significance among group (*$P < 0.05$) was assessed by Welch's *t* test.

(Fig. 4, B to E). Under resting-NREM sleep for 30 min in total, M2 and S1 were synchronously activated at 2 Hz with an in-phase (Fig. 4, B and C) or an antiphase pattern (Fig. 4, D and E). Synchronous activation did not change task performance in the testing period 1 day after the sampling period. In wild-type mice, novelty preference decayed after 2 days (fig. S3), but synchronous stimulation prolonged memory retention to at least 4 days after the sampling period (Fig. 4F). In contrast, antisynchronous activation resulted in a decrease in performance in the testing period 1 day after the sampling period (Fig. 4F). Synchronized coactivation of M2 and S1 during SD promoted memory consolidation over 4-day intervals (Fig. 4, G and H).

In this study, we showed that perceptual memory consolidation requires top-down cortical input during NREM sleep. Memory consolidation was dependent on causal frontoparietal information flow during ~4% of total sleep time (inactivation during the cumulative 30 min of NREM sleep versus the 12 hours of total sleep in mice) (*22*). Our findings demonstrate a causal relationship between cortical top-down projections and reactivated neurons for memory consolidation and suggest a general hierarchical



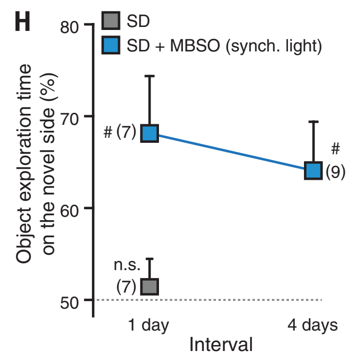


Fig. 4. Memory consolidation depends on the phase synchrony of slow waves. (**A**) Optogenetic activation of M2 and S1 with LFP recordings with single electrodes. (**B**) LFPs and multiunit activity (MUA) recordings from M2 (top) and S1 (bottom) during resting-NREM sleep with synchronous photostimulation. (**C**) Normalized firing activities during light on (normalized to interval off periods). (**D**) LFPs and MUA with antisynchronous photostimulation. (**E**) Normalized firing activities. (**F**) Task performance after photostimulation during resting-NREM sleep. (**G**) (Left) Behavioral paradigm. After the sampling period, transgenic mice were deprived of sleep for 1 hour with synchronized coactivation of M2 and S1. (Right) Photostimulation protocol. The photostimulation (2 Hz) during the SD experiment was applied according to the NREM sleep pattern from mouse A. (**H**) Summary of task performance after photostimulation. The cumulative illumination time was 30 min. Statistical significance from off period firing rate (#P < 0.05, ##P < 0.01) or from 50% chance level (#P < 0.05) was assessed by one-sample t test.

control by presynaptic neurons in higher cortical areas to regulate memory consolidation in lower sensory areas (*23*). Reactivation of cortico-cortical regions underlying these top-down circuits could enhance memory retention periods (Fig. 4F). Further, synchronized coactivation of M2 and S1 could overcome the physiologically adverse effects of SD and retain a long-term memory (Fig. 4H). These results indicate that perceptual memory consolidation may not require sleep per se, but that synchronized coactivation of hierarchical cortical pathways, enabled by SWA in NREM sleep, is necessary and sufficient for sensory experience to be consolidated in memory.

REFERENCES AND NOTES

1. G. Yang *et al.*, *Science* **344**, 1173–1178 (2014).
2. A. Rolls *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13305–13310 (2011).
3. M. Massimini, R. Huber, F. Ferrarelli, S. Hill, G. Tononi, *J. Neurosci.* **24**, 6862–6870 (2004).
4. Y. Nir *et al.*, *Neuron* **70**, 153–169 (2011).
5. V. V. Vyazovskiy *et al.*, *Nature* **472**, 443–447 (2011).
6. G. Tononi, C. Cirelli, *Neuron* **81**, 12–34 (2014).
7. M. Inostroza, J. Born, *Annu. Rev. Neurosci.* **36**, 79–102 (2013).
8. S. Diekelmann, J. Born, *Nat. Rev. Neurosci.* **11**, 114–126 (2010).
9. S. Manita *et al.*, *Neuron* **86**, 1304–1316 (2015).
10. R. A. Bevins, J. Besheer, *Nat. Protoc.* **1**, 1306–1311 (2006).
11. S. L. Dix, J. P. Aggleton, *Behav. Brain Res.* **99**, 191–200 (1999).
12. M. Inostroza, S. Binder, J. Born, *Behav. Brain Res.* **237**, 15–22 (2013).
13. M. N. de Lima, D. C. Laranja, E. Bromberg, R. Roesler, N. Schröder, *Behav. Brain Res.* **156**, 139–143 (2005).
14. J. Chen, A. R. Kriegstein, *Science* **350**, 554–558 (2015).
15. H. Zeng *et al.*, *Cell* **107**, 617–629 (2001).
16. T. Nakashiba, J. Z. Young, T. J. McHugh, D. L. Buhl, S. Tonegawa, *Science* **319**, 1260–1264 (2008).
17. K. G. Phillips *et al.*, *Neuron* **76**, 526–533 (2012).
18. S. Ribeiro *et al.*, *Front. Neurosci.* **1**, 43–55 (2007).
19. A. Brovelli *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9849–9854 (2004).
20. M. Kamiński, M. Ding, W. A. Truccolo, S. L. Bressler, *Biol. Cybern.* **85**, 145–157 (2001).
21. R. Beltramo *et al.*, *Nat. Neurosci.* **16**, 227–234 (2013).
22. J. L. Valatx, R. Bugat, *Brain Res.* **69**, 315–330 (1974).
23. Y. Dudai, A. Karni, J. Born, *Neuron* **88**, 20–32 (2015).

ACKNOWLEDGMENTS

We thank C. Yokoyama for helpful comments and editing, H. Hirase, T. Toyoizumi, S. Fujisawa, J. P. Johansen and B. Mensh for comments and discussions, S. Itohara and K. Yasuda for behavioral analysis, T. Tsunematsu for sleep measurement, M. Kobayashi and H. Hioki for anatomical analysis, R. Endo and R. Kato for animal control, K. Tao for a part of firing data analysis, and A. Kamoshida for writing custom-made LabVIEW programs. The archaerhodopsin from the Halorubrum strain TP009 (AAV-CMV-ArchT-EGFP) was kindly given by E. Boyden (MIT, Cambridge, MA, USA). The authors acknowledge support from a Grant-in-Aid for Young Scientists (A) from the Japan Society for the Promotion of Science to M.M.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1315/suppl/DC1
Materials and Methods
Figs. S1 to S18
References (*24–40*)

25 December 2015; accepted 12 May 2016
Published online 26 May 2016
10.1126/science.aaf0902

NEURONAL PLASTICITY

Cell-specific restoration of stimulus preference after monocular deprivation in the visual cortex

Tobias Rose,* Juliane Jaepel, Mark Hübener, Tobias Bonhoeffer*

Monocular deprivation evokes a prominent shift of neuronal responses in the visual cortex toward the open eye, accompanied by functional and structural synaptic rearrangements. This shift is reversible, but it is unknown whether the recovery happens at the level of individual neurons or whether it reflects a population effect. We used ratiometric Ca^{2+} imaging to follow the activity of the same excitatory layer 2/3 neurons in the mouse visual cortex over months during repeated episodes of ocular dominance (OD) plasticity. We observed robust shifts toward the open eye in most neurons. Nevertheless, these cells faithfully returned to their pre-deprivation OD during binocular recovery. Moreover, the initial network correlation structure was largely recovered, suggesting that functional connectivity may be regained despite prominent experience-dependent plasticity.

How do mature cortical circuits achieve a stable representation of sensory input while maintaining their capacity to adapt to changes in an animal's sensory environment? In the binocular visual cortex, afferent inputs from the two eyes converge onto single neurons. Monocular deprivation (MD) shifts ocular dominance (OD) toward the nondeprived eye. This OD shift can be induced in juvenile carnivorans (*1*, *2*) and primates (*3*), as well as in rodents (*4*–*10*). In mice, OD shifts also occur in adults and are reversible (*7*, *11*–*13*), providing the opportunity for longer-lasting experiments and the study of consecutive phases of experience-dependent plasticity at the single-cell level. Due to technical limitations, however, longitudinal assessments of OD plasticity have so far been based on either short-term single-cell recordings (*14*), recordings from different animals (*15*, *16*), or repeated population recordings lacking single-cell resolution (*7*, *11*–*13*, *17*). It is therefore not yet clear how individual neurons shift their functional properties in response to prolonged perturbations of sensory input, and whether their initial stimulus selectivity is lost or maintained after recovery. In this study, we measured the changes in the response properties of single neurons during MD, recovery from MD, and repeated MD.

We performed ratiometric Ca^{2+} imaging (fig. S1), using the genetically encoded Ca^{2+} indicator GCaMP6s (*18*), which we coexpressed with the bright structural marker mRuby2 (*19*) under control of the CamKII0.4 promoter, using adeno-associated virus. This allowed us to follow visually evoked activity of the same excitatory (fig. S2) layer 2/3 neurons in the binocular cortex of adult mice over months during multiple episodes of OD plasticity (Fig. 1A).

A substantial fraction (64%) of excitatory layer 2/3 neurons responded to eye-specific drifting grating stimuli of varying orientation with significant changes in their somatic Ca^{2+} concentration (Figs. 1, B and C, and 2C). We closed the lid of the eye contralateral to the recorded hemisphere for 5 to 8 days under conditions that facilitate adult OD plasticity (see supplementary methods). The majority (62%) of all continuously responsive cells showed significant OD changes (Fig. 2, B and C), and most cells shifted toward the nondeprived eye. Contrary to what classical theories would predict, ~10 to 20% of the cells responded more strongly to deprived-eye stimulation after MD (Fig. 2B, fig. S3, and supplementary text).

The overall number of visually responsive cells decreased significantly after MD (Fig. 2C). As in previous reports using a similar deprivation paradigm in adult mice (*8*, *9*), we observed strong deprived-eye depression and more moderate open-eye strengthening (Fig. 2D). Unlike OD plasticity in juvenile mice (*10*), a major driving force was deprived-eye depression of initially contralaterally dominated neurons lacking prominent competing open-eye input (Fig. 2D). Whereas the sum of the eye-specific changes in response strength remained largely constant in originally binocular neurons (Fig. 2D), cells with a strong contralateral bias were prone to be silenced during MD, explaining the decrease in responsive neurons.

OD plasticity is based on substantial modifications to visual cortex circuits, including changes in excitatory and inhibitory synapse number and strength (*4*, *5*, *13*, *20*–*22*). Deprived-eye depression and open-eye strengthening leave deprived-eye inputs at a competitive disadvantage during binocular recovery. Both factors may render it unlikely that eye-specific responses recover precisely to their original values in single cells. Instead, recovery may be mediated by compensatory changes at the network level, irrespective of the initial selectivity of individual neurons. Alternatively, some "memory trace" of the initial eye-specific input strength may guide individual neurons to return to their original response properties after MD.

We allowed 8 to 10 days of recovery with normal binocular vision and assessed the eye-specific responsiveness of significantly plastic cells (MD shift magnitude >1σ of baseline OD fluctuations). The population of plastic neurons recovered fully (Fig. 2E). Compensatory changes of neurons other than the plastic cells are therefore unnecessary to explain recovery from MD. To assess the degree to which individual neurons regain their initial OD after MD, we compared the OD after recovery with the pre-MD OD of plastic neurons. Single-cell OD just before MD and after recovery were significantly correlated [Pearson's correlation coefficient (r) = 0.51], although individual neurons showed deviations from the unity line of perfect recovery (Fig. 2F and fig. S4A).

To gauge the magnitude of MD-induced changes and the degree of single-cell recovery, we quantified the variability of eye-specific response properties under baseline conditions (Fig. 2, G and H, and fig. S5). Population OD was stable over ~8 days of baseline recording (three imaging sessions) and showed a characteristic contralateral bias (fig. S5B). To assess the stability of singe-cell tuning, we measured the functional properties of the same repeatedly imaged neurons over different time intervals (4 days and 12 to 14 days). If neurons in the binocular visual cortex would continuously and randomly change their response properties in a process akin to a random walk, one would expect drift, in which net changes in response properties accumulate over time. However, we found no systematic difference between the changes in OD, orientation preference, and orientation selectivity for the short and long measuring interval (Fig. 2H and fig. S5, C to H), indicating the presence of mechanisms ensuring the constancy of neural responses over time.

Still, baseline variability was considerably higher than expected from measurement uncertainty alone (Fig. 2H, within-session correlation). Thus, similar to but less pronounced than in the somatosensory cortex (*23*), neurons in the visual cortex show response variation even during normal sensory experience. The difference may be related to the different rates of synaptic turnover in these sensory areas (*4*, *24*). For the same cells, baseline variability was higher in awake mice than under anesthesia (fig. S7).

The single-cell OD index (ODI) deviated prominently from baseline after MD (Fig. 2H and fig. S4). The baseline-recovery correlation of individual plastic neurons, however, was not different from the 4-day within-baseline correlation (Fig. 2H), and the distribution of pairwise ODI differences was equally indiscernible from the baseline (fig. S4). Orientation tuning was largely unaffected by MD (fig. S6).

Having shown that the OD of individual cells is restored to its original value within the bounds set by baseline variability, we next asked whether the underlying eye-specific inputs also regain their original strengths. On the population level, ipsi- and contralateral response magnitudes were largely reestablished after recovery (Fig. 3A). To quantify the degree to which the combined eye-specific changes during recovery mirror the changes

Max Planck Institute of Neurobiology, Am Klopferspitz 18, D-82152 Martinsried, Germany.
*Corresponding author. Email: trose@neuro.mpg.de (T.R.); tobias.bonhoeffer@neuro.mpg.de (T.B.)

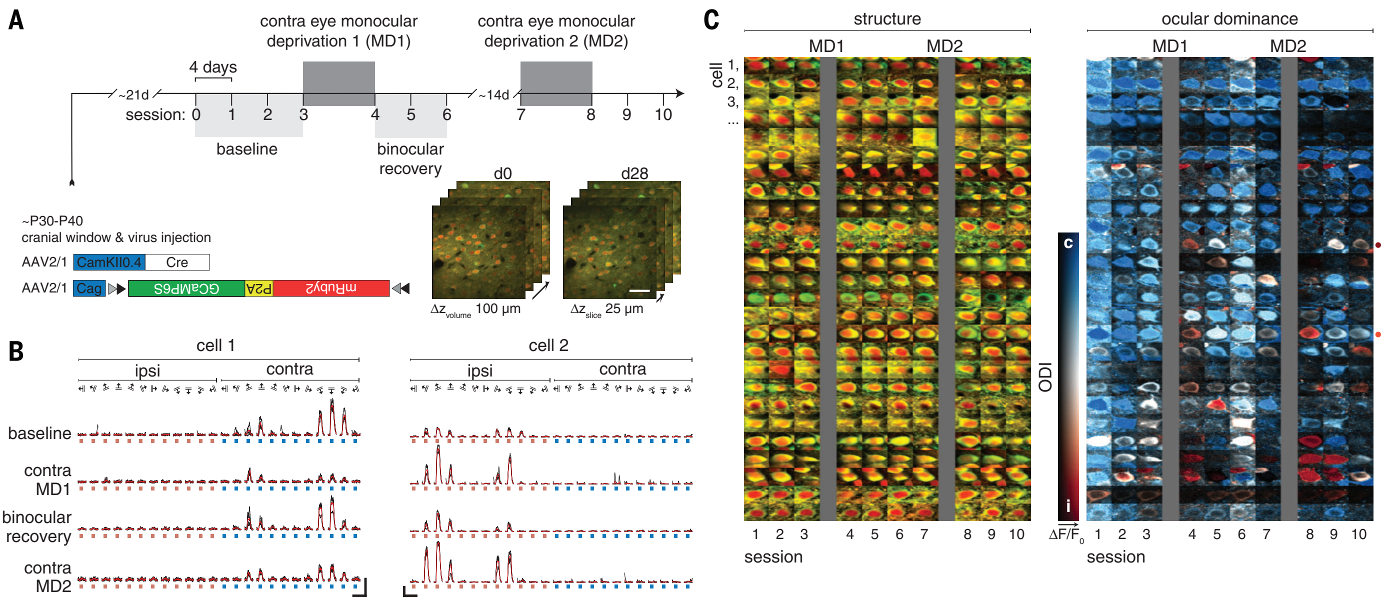


Fig. 1. Ratiometric long-term single-cell imaging of repeated OD plasticity. (**A**) Timeline of the experiment. We performed chronic ratiometric Ca^{2+} imaging (mRuby2-P2A-GCaMP6s) of the somata of excitatory layer 2/3 neurons in a small volume of the binocular visual cortex in adult mice [rapid sequential acquisition of a 185 × 185 × 100 μm volume; four slices at an image plane depth increment (Δz_{slice}) of 25 μm, frame rate 7.5 Hz, scale bar 40 μm]. The same image locations and cells were revisited for up to 2 months during multiple episodes of OD plasticity [contralateral eye MD (contra MD), binocular recovery, and repeated contra MD]. (**B**) Ca^{2+} signals (fluorescence ratio changes, $\Delta R/R_0$) in two example neurons in response to ipsi- and contralateral eye drifting grating stimulation (12 directions, four repetitions) over selected sessions covering all OD plasticity episodes. After each of two MD periods, a contralaterally (cell 1) and an ipsilaterally (cell 2) dominated cell exhibit repeated deprived-eye depression or open-eye strengthening, respectively, and show response recovery during binocular vision in between (scale bars $\Delta R/R_0$ = 200%, 10 s; for further examples, see figs. S5A, S6B, and S10C). (**C**) Sorted structural (left) and functional (right) cell maps of individual neurons imaged over 10 sessions (26 continuously responsive cells of a single animal showing a significant OD shift during the first MD). OD is depicted as the pixel-wise peak fluorescence ratio change in response to ipsi- and contralateral eye preferred grating direction [$ODI = (\Delta R/R_{contra} - \Delta R/R_{ipsi})/(\Delta R/R_{contra} + \Delta R/R_{ipsi})$]. Red hues indicate ipsilateral dominance (ODI < 0), and blue hues indicate contralateral dominance (ODI > 0). Pixel intensity is scaled by response amplitude. Cells are sorted by cell identity (horizontally) and by ODI of the last pre-MD session (vertically, descending ODI; red dots: example neurons shown in Fig. 3, D and E).

during MD on the single-cell level, we expressed the change in ipsi- and contralateral single-cell responsiveness during MD and recovery as vectors, and used the angular difference between these shift trajectories as a metric for the similarity between initial shift and subsequent recovery (Fig. 3B). Recovery of OD in individual plastic neurons was achieved by highly specific changes in ipsi- and contralateral responsiveness that reversed the eye-specific changes induced by the first MD on the single-cell level (Fig. 3C). The precision of single-cell recovery was significantly greater than would be expected if neurons simply followed the global population change of eye-specific input strength (Fig. 3C, shuffled data). These findings, including precise ODI recovery, could be reproduced in the awake mouse (fig. S8).

It has been proposed that the synaptic connections that have been established during an early MD episode are reused during a second MD (*4*, *11*). To test whether the same neurons undergo repeated plasticity and to assess whether these cells repeat the previously observed eye-specific changes, we performed a second MD 2 weeks after the last recovery session (Fig. 3, D to F). Single-cell OD changes after the second MD correlated significantly with the OD changes induced by the first MD in the same neurons (Fig. 3E). The relative contributions of ipsi- and contralateral response changes during the first and second MD were, again, highly correlated on the single-cell level (Fig. 3F) and could not simply be explained by a population-wide effect (Fig. 3F, shuffled data). Therefore, individual neurons undergo repeated plastic changes, with highly reproducible bidirectional modifications in eye-specific inputs, over repeated periods of experience-dependent plasticity (MD, recovery, second MD).

Both MD and recovery lead to prominent structural and functional synaptic rearrangements (*4*, *5*, *13*, *20*–*22*), which are indicative of major network rewiring. This raises the question of to what extent not only single-cell responses but also connectivity among cells are recovered after such phases of dramatic functional plasticity. We analyzed the structure of pairwise correlations in trial-to-trial fluctuations (i.e., noise correlations) and mean stimulus responses (i.e., signal correlations) over time (Fig. 4). Pairwise noise correlations in spike trains are often used as a proxy for functional connectivity (*25*, *26*). Pairwise signal correlations, in turn, measure evoked response similarity and are a more comprehensive measure of neuronal feature selectivity than are unidimensional tuning indices such as ODI (*27*). Pairs of neurons showing high signal correlation have a high probability of being directly connected (*28*).

We found a broad distribution of pairwise noise and signal correlations with low average correlation values (fig. S9, A and D). To assess the stability of network activity patterns, we followed the similarity of correlation structures throughout the plasticity episodes. We correlated the matrices of pairwise correlation coefficients with each other (*29*) to obtain a similarity measure (Fig. 4A). MD induced a drop in both signal and noise correlation matrix similarity in comparison to the last baseline session. After MD signal and noise correlation similarity returned to original values (Fig. 4B). Thus, specific network activity patterns are altered by MD, but recovery from MD renders them indistinguishable from those before MD. Therefore, the correlation structure of neuronal activity patterns, and thus perhaps functional connectivity, seem to be remarkably resilient to massive perturbations of sensory input.

One explanation for this remarkable stability in the mature visual cortex may be that a subset of stable synaptic connections is protected from being overwritten by plasticity, thereby providing the "tuning backbone" that on one hand counteracts the degradation or drift of eye-specific responses during baseline and on the other hand guides precise recovery after major plasticity episodes (Fig. 2H). Indeed, recent data suggest that cells with correlated responses form functional subnetworks

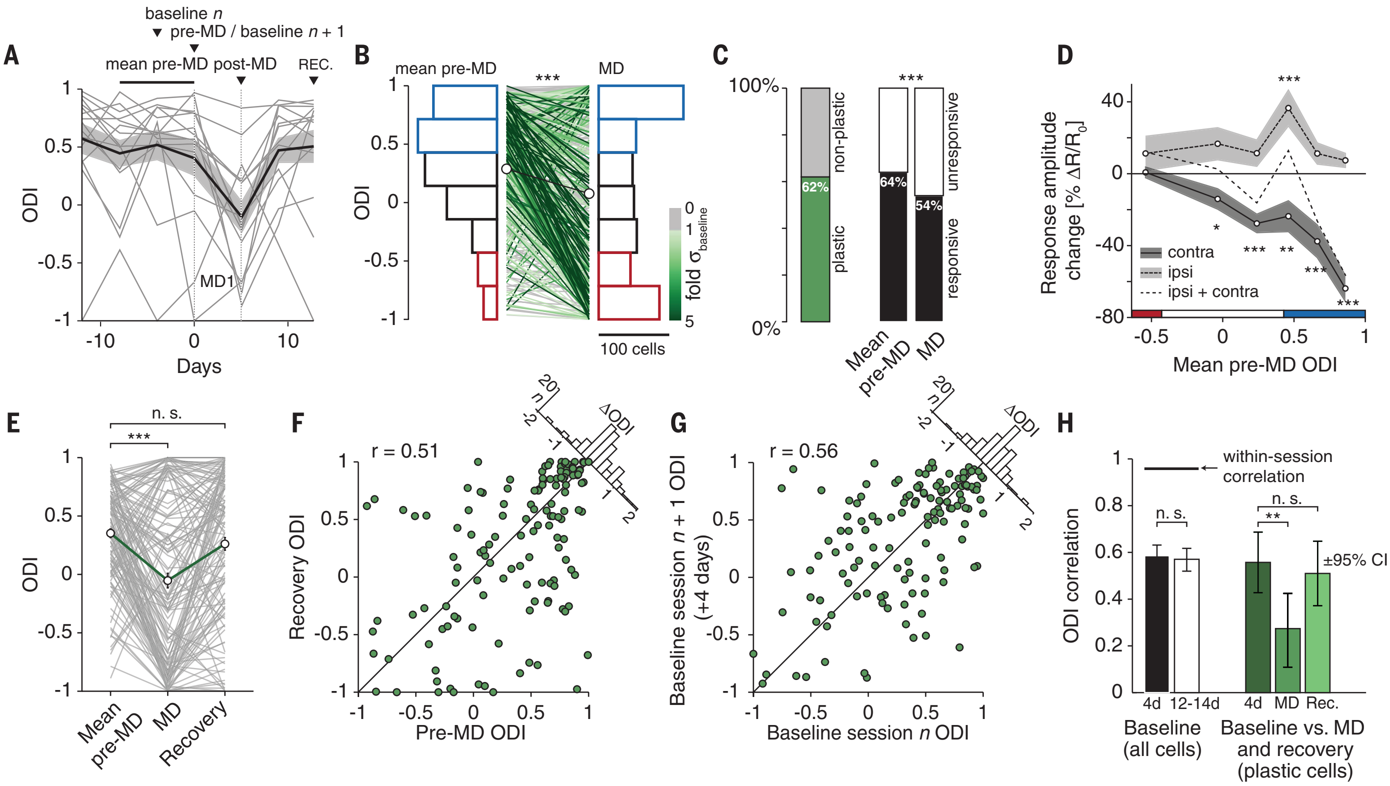


Fig. 2. Full recovery of eye-specific tuning after MD. (**A**) Example time courses of single-cell ODI over baseline, contralateral eye MD, and recovery (gray lines: continuously responsive individual cells; black line: mean ± SEM, $n = 15$ cells, 1 animal). (**B**) Single-cell ODI distribution during baseline (mean of three pre-MD sessions: ODI = 0.29 ± 0.47, SD) and after 5 to 8 days of contralateral eye MD (post-MD: ODI = 0.08 ± 0.67, SD, $n = 456$ cells, 10 animals, $P < 10^{-12}$, Wilcoxon signed-rank test). The lines connect individual neurons; line color indicates shift significance expressed in units of standard deviations over baseline fluctuations. Colored histogram bins indicate class definitions for contralateral (blue), binocular (black), and ipsilateral (red) cells. (**C**) Left, fraction of neurons showing a significant MD-evoked ODI change ($n = 456$ cells). Right, fraction of responsive and unresponsive neurons before and after MD ($P < 10^{-6}$, $n = 1245$ cells, χ^2 test). (**D**) Difference in eye-specific responsiveness before and after MD (gray shading: ±SEM) as a function of pre-MD ODI [$n = 593$ cells continuously responsive during three baseline sessions, grouped into pre-MD ODI sextiles comprising similar numbers of cells (98 to 99 cells per class), paired t tests; OD group classes are indicated by the color bar on the x axis]. (**E**) Population ODI recovery of plastic neurons (ODI change >1σ of baseline fluctuations; mean ± SEM; $n = 133$ cells from 8 animals continuously responsive during baseline, MD, and recovery; baseline versus MD $P < 10^{-7}$, baseline versus recovery $P > 0.09$, Wilcoxon signed-rank test; gray lines: individual cells). (**F**) Single-cell ODI correlation of significantly plastic neurons before MD and after recovery (pre-MD versus recovery Pearson's $r = 0.51$, $P < 10^{-9}$, $n = 133$ cells, 8 animals). (**G**) Correlation of the same cells during baseline sessions spaced 4 days apart ($r = 0.56$, $P < 10^{-11}$). Histograms in (F) and (G) show the distributions of pairwise differences (not drawn to scale; not significantly different, $P = 0.35$, two-sample Kolmogorov-Smirnov test, $n = 133$ cells, 8 animals; fig. S4). (**H**) Left, baseline 4-day and 12- to 14-day eye-specific visual tuning correlations of all responsive neurons ($n = 738$ cells, 11 animals, 4-day versus 12- to 14-day $P > 0.65$, Fisher's r-to-z transformation). To estimate the influence of trial-to-trial variability, we show the distribution of bootstrapped within-session correlations (black line). Right, correlation of the same plastic neurons during baseline (sessions 4 days apart), baseline and MD (fig. S4), and baseline and recovery ($n = 133$ cells, 4-day versus MD $P < 0.006$; 4-day versus recovery (Rec.) $P > 0.59$, Fisher's r-to-z transformation). Correlations are presented with 95% confidence intervals (CIs). Here and in the following figures, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

that are interconnected by exceptionally strong synapses (*28*). It is tempting to speculate that these connections remain stable even under conditions of a changing sensory environment, whereas the majority of weak synapses provide the substrate for reversible plastic modifications (*28*, *30*). It remains to be established whether much longer deprivation episodes would lead to less precise recovery, especially when initiated early in life, when the rearrangements of thalamocortical projections have been reported to be extensive (*31*).

Our data clearly show that experience-dependent plasticity of neurons in the visual cortex occurs on a cell-by-cell basis. Many of the experience-dependent changes are carried by a subpopulation of plastic neurons, which regain their original response properties after plasticity is reversed and follow the same eye-specific shift trajectories when challenged with a second deprivation episode. However, many neurons are resilient to plastic changes and some even change their response properties in the opposite direction of what classical theories would have predicted. The role of the counterintuitive shifters is less clear, but neurons resilient to plastic changes could be important for providing stability of the cortical network in the face of constantly occurring changes and adaptations. Neurons that change strongly but recover precisely, together with neurons not susceptible to plastic changes, may be the visual cortex's solution to the opposing needs for plasticity and stability.

REFERENCES AND NOTES

1. T. N. Wiesel, D. H. Hubel, *J. Neurophysiol.* **26**, 1003–1017 (1963).
2. A. Kirkwood, M. C. Rioult, M. F. Bear, *Nature* **381**, 526–528 (1996).
3. S. Le Vay, T. N. Wiesel, D. H. Hubel, *J. Comp. Neurol.* **191**, 1–51 (1980).
4. S. B. Hofer, T. D. Mrsic-Flogel, T. Bonhoeffer, M. Hübener, *Nature* **457**, 313–317 (2009).
5. S. J. Kuhlman *et al.*, *Nature* **501**, 543–546 (2013).
6. J. A. Gordon, M. P. Stryker, *J. Neurosci.* **16**, 3274–3286 (1996).
7. N. B. Sawtell *et al.*, *Neuron* **38**, 977–985 (2003).
8. U. Matthies, J. Balog, K. Lehmann, *J. Neurosci.* **33**, 11774–11778 (2013).
9. F. Greifzu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1150–1155 (2014).
10. T. D. Mrsic-Flogel *et al.*, *Neuron* **54**, 961–972 (2007).
11. S. B. Hofer, T. D. Mrsic-Flogel, T. Bonhoeffer, M. Hübener, *Nat. Neurosci.* **9**, 127–132 (2006).
12. M. Kaneko, J. L. Hanover, P. M. England, M. P. Stryker, *Nat. Neurosci.* **11**, 497–504 (2008).
13. D. van Versendaal *et al.*, *Neuron* **74**, 374–383 (2012).

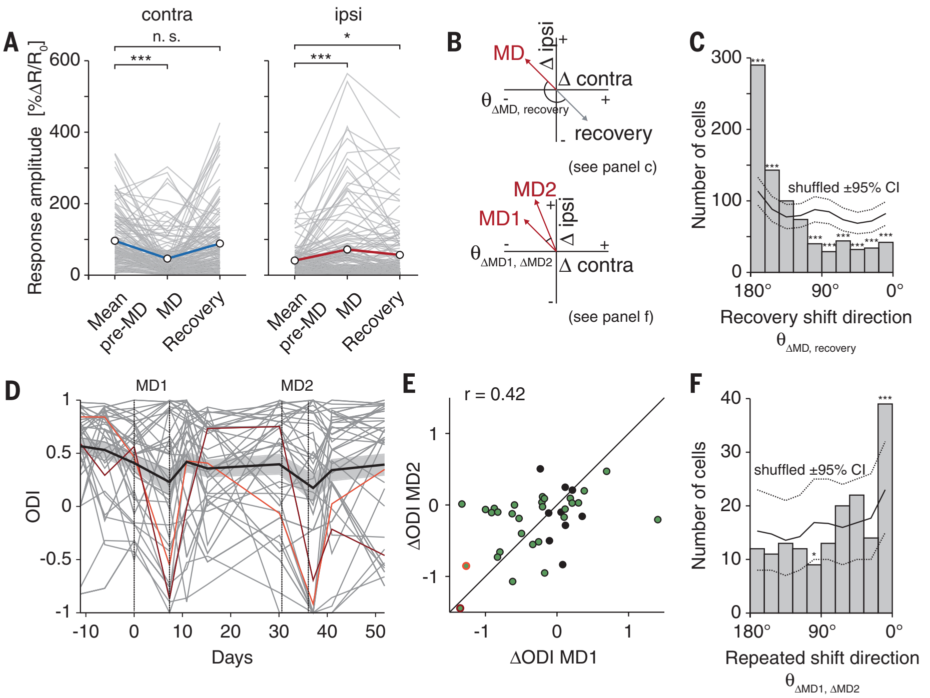


Fig. 3. Precise reversion and repetition of eye-specific changes during recovery and repeated MD. (**A**) Population recovery of eye-specific response amplitudes of plastic neurons (mean ± SEM; n = 133 cells, 8 animals; contra: baseline versus MD $P < 10^{-10}$, baseline versus recovery $P > 0.3$; ipsi: baseline versus MD $P < 0.0002$, baseline versus recovery $P < 0.02$; gray lines: individual cells; paired t test). (**B**) Schematics illustrate the calculation of the difference between the directions of initial shift and recovery (upper panel) or between shifts induced by the first and second MD (lower panel; see supplementary methods). (**C**) Single-cell angular differences ($\theta_{\Delta MD,recovery}$) between MD and recovery shift trajectories. The difference of 180° indicates perfectly sign-inverted shift in joint ipsi- and contralateral responsiveness after recovery (gray bars: angular differences, line: bootstrapped angular differences with shuffled cell IDs between MD and recovery sessions ± 95% CI; n = 828 cells). (**D**) Example time course of single-cell ODI over baseline, MD, recovery, and repeated contra MD [gray lines: individual, continuously responsive cells; black line: mean ± SEM, n = 36 cells, 1 animal; dark red and light red lines indicate example neurons shown in (E) and in Fig. 1C (red dots)]. (**E**) Correlation of single-cell ODI shift (post-MD − pre-MD ODI) between first and second MD (Pearson's r = 0.42, P = 0.009, n = 38 cells, 2 animals; green symbols: cells significantly plastic during first MD). (**F**) Single-cell angular differences ($\theta_{\Delta MD1,\Delta MD2}$) between MD1 and MD2 shift trajectories. A difference of 0° indicates identical shifts in joint ipsi- and contralateral responsiveness during first and second MD (line: bootstrapped angular differences with shuffled cell IDs between MD1 and MD2 sessions ± 95% CI; n = 165 cells). Significance in (C) and (F) is assessed by comparing matched cell IDs against the 95, 99, and 99.9% CIs of bootstrapped angular differences from scrambled cell IDs.

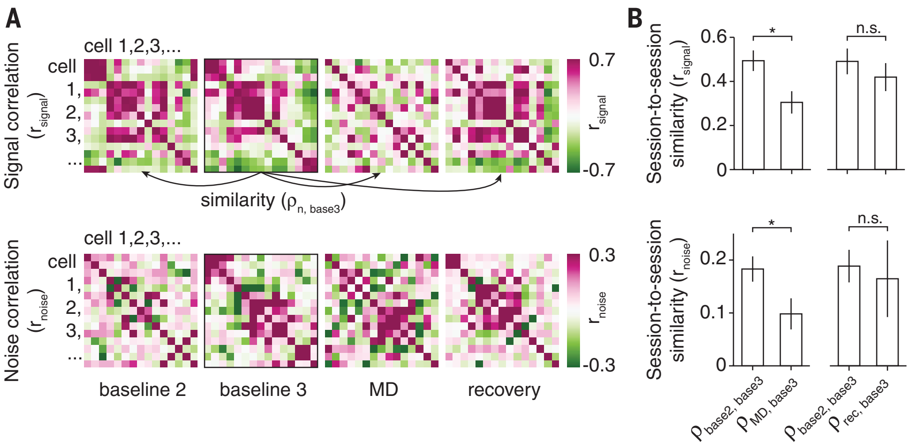


Fig. 4. Recovery of correlation structure. (**A**) Example pairwise signal (r_{signal}) and noise (r_{noise}) correlation structure of 15 consistently responsive neurons of a single animal imaged over baseline, MD, and recovery sessions. For visualization of correlation structure only, cells are clustered and sorted independently for r_{signal} and r_{noise} on the last pre-MD session (supplementary methods). The same sort matrix is used for all other time points. As a measure of correlation structure similarity, we used Pearson's r of the correlation matrices with baseline session 3 ($\rho_{n,3}$). (**B**) Average session-wise similarity of r_{signal} and r_{noise} structures with the last baseline session (mean ± SEM). Both signal and noise correlation structure similarity show a drop after MD (r_{signal}: P = 0.011, r_{noise}: P = 0.011, n = 9 animals, Wilcoxon signed-rank test), but after recovery both r_{signal} and r_{noise} correlation structures are not statistically different from baseline (r_{signal}: P = 0.11, r_{noise}: P = 0.38, n = 7 animals).

14. L. Mioche, W. Singer, *J. Neurophysiol.* **62**, 185–197 (1989).
15. T. E. Krahe, A. E. Medina, R. E. de Bittencourt-Navarrete, R. J. Colello, A. S. Ramoa, *Neuron* **48**, 329–343 (2005).
16. T. N. Wiesel, D. H. Hubel, *J. Neurophysiol.* **28**, 1060–1072 (1965).
17. D. S. Kim, T. Bonhoeffer, *Nature* **370**, 370–372 (1994).
18. T.-W. Chen *et al.*, *Nature* **499**, 295–300 (2013).
19. A. J. Lam *et al.*, *Nat. Methods* **9**, 1005–1012 (2012).
20. N. Mataga, Y. Mizuguchi, T. K. Hensch, *Neuron* **44**, 1031–1041 (2004).
21. J. L. Chen *et al.*, *Neuron* **74**, 361–373 (2012).
22. K. L. Villa *et al.*, *Neuron* **89**, 756–769 (2016).
23. D. J. Margolis *et al.*, *Nat. Neurosci.* **15**, 1539–1546 (2012).
24. A. J. G. D. Holtmaat *et al.*, *Neuron* **45**, 279–291 (2005).
25. B. Doiron, A. Litwin-Kumar, R. Rosenbaum, G. K. Ocker, K. Josić, *Nat. Neurosci.* **19**, 383–393 (2016).
26. M. N. Shadlen, W. T. Newsome, *J. Neurosci.* **18**, 3870–3896 (1998).
27. G. Rothschild, I. Nelken, A. Mizrahi, *Nat. Neurosci.* **13**, 353–360 (2010).
28. L. Cossell *et al.*, *Nature* **518**, 399–403 (2015).
29. A. Luczak, P. Barthó, K. D. Harris, *Neuron* **62**, 413–425 (2009).
30. S. Fusi, P. J. Drew, L. F. Abbott, *Neuron* **45**, 599–611 (2005).
31. A. Antonini, M. Fagiolini, M. P. Stryker, *J. Neurosci.* **19**, 4388–4406 (1999).

ACKNOWLEDGMENTS

We thank T. Mrsic-Flogel for discussions, P. Goltstein for discussions and programming help, and C. Huber and V. Staiger for help with molecular cloning and immunohistochemistry. This work was supported by the Max Planck Society, a Marie Curie Intra-European Fellowship to T.R., a Boehringer Ingelheim Ph.D. fellowship to J.J., and the Deutsche Forschungsgemeinschaft (grant SFB870 to T.R. and T.B.). Data and analysis codes are stored and curated at the Max Planck Computing and Data Facility Garching, Munich. Plasmids have been deposited at Addgene.org (accession no. 68717).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1319/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
References (*32–52*)

31 August 2015; accepted 12 May 2016
10.1126/science.aad3358

BRAIN RESEARCH

Prospective representation of navigational goals in the human hippocampus

Thackery I. Brown,[1]* Valerie A. Carr,[1,2] Karen F. LaRocque,[1] Serra E. Favila,[3] Alan M. Gordon,[1] Ben Bowles,[4] Jeremy N. Bailenson,[5] Anthony D. Wagner[1,6]*

Mental representation of the future is a fundamental component of goal-directed behavior. Computational and animal models highlight prospective spatial coding in the hippocampus, mediated by interactions with the prefrontal cortex, as a putative mechanism for simulating future events. Using whole-brain high-resolution functional magnetic resonance imaging and multi-voxel pattern classification, we tested whether the human hippocampus and interrelated cortical structures support prospective representation of navigational goals. Results demonstrated that hippocampal activity patterns code for future goals to which participants subsequently navigate, as well as for intervening locations along the route, consistent with trajectory-specific simulation. The strength of hippocampal goal representations covaried with goal-related coding in the prefrontal, medial temporal, and medial parietal cortex. Collectively, these data indicate that a hippocampal-cortical network supports prospective simulation of navigational events during goal-directed planning.

Prospective thought and the simulation of future experiences are fundamental for planning how to best achieve immediate and longer-term goals. Prospection is theorized to rely on neural mechanisms that underlie episodic memory (*1*, *2*), drawing on declarative memory for distinct events to flexibly simulate future experiences and outcomes. The hippocampus subserves episodic retrieval of goal-relevant spatial sequences in rodents (*3*–*7*) and humans (*8*–*12*) and plays a central role in models of goal-directed navigation and episodic memory (*13*–*15*). In rodents, hippocampal "place cells" exhibit prospective sequential firing along navigational routes during planning that reflects current goals (*16*, *17*). Prospective firing may support reinstatement of the multifeatural representations of spatial contexts in a broader network underlying prospection and goal coding [including the medial temporal lobe (MTL), retrosplenial complex (RSC), and ventral striatum (VS)] (*1*, *2*, *18*–*21*). Prospective simulation may also rely on hippocampal interactions with the prefrontal cortex (PFC), which may provide cognitive control machinery through which mnemonic details are flexibly accessed and combined into the formulation of future route plans (*22*, *23*). A fundamental question in human cognitive neuroscience is whether the hippocampus and its functional interactions support flexible prospective representation of spatial trajectories during goal-directed planning.

Although human hippocampal neurons demonstrate location- and goal-related responses that can be reinstated during retrieval (*24*, *25*), noninvasive quantification of the neural representation of spatial information in humans is a challenge. Functional magnetic resonance imaging (fMRI) has revealed distance-to-goal (*26*–*28*) and grid cell–like (*29*) response coding in the human hippocampus and entorhinal cortex. Measurement of purely place cell–based location codes may not be feasible with fMRI; however, it may be possible to quantify episodic retrieval of a distributed multifeatural engram of a spatial context. Multivariate fMRI approaches have demonstrated that distributed patterns in the hippocampus, MTL cortex, and RSC carry representational information about environmental features, locations, and the direction to a goal (*30*–*34*). However, direct evidence that this hippocampal-cortical network supports prospective goal coding during route planning in humans has yet to be shown.

We used whole-brain high-resolution fMRI (hr-fMRI; 1.6-mm isotropic voxels) to simultaneously record fine-grained pattern information from the human hippocampus and a core network of anatomically and functionally interconnected regions putatively involved in goal coding and prospection (supplementary materials). Participants underwent hr-fMRI while performing a virtual navigation paradigm designed to parallel tasks that have been used with rodents (*17*, *35*). On day 1, outside the scanner, participants learned to navigate to five goal locations in a virtual circular environment, each marked by a distinct pair of fractal images (Fig. 1, A and B). On day 2, while undergoing hr-fMRI, participants began each trial at one of the locations; their viewpoint then shifted toward the ground, and they were cued with one of the fractals to plan navigation of the shortest route from their current position to the cued goal location (planning period). The participant's view then panned up, and they actively navigated to the goal. Critically, fractals were no longer visible at the goal locations on day 2, and thus performance depended on memory (Fig. 1C). During scanning, participants planned and executed navigation between the five locations across 160 trials (32 per location, visiting every location from every other location an equal number of times). This design enabled analysis of neural patterns during planning that represent information about future goal states—information that generalizes across cues, start positions, and routes.

We used multi-voxel pattern analyses to classify planning period activity (before active navigation) as being related to the current location ("current" classifier) or the future goal location to which participants would navigate ("future" classifier). We quantified current-state and future goal–state representations and their relative strength on a region-by-region and trial-by-trial basis by using classifier accuracy (significance measured against empirically validated chance; supplementary materials) and probabilistic evidence scores. In hypothesis-driven analyses, we analyzed data from a priori anatomical regions of interest (ROIs). We indexed the representation of navigational events within the hippocampus and examined how hippocampal representations covary with (i) goal-related codes in the MTL cortex, RSC, and VS and (ii) planning activity in the PFC.

On day 2, participants were highly accurate at cued navigation, performing near ceiling levels (supplementary materials). Applying the "current" classifier to the planning period data, we confirmed that distributed patterns of human hippocampal activity code for current location (classifier accuracy, 29.9%; $t_{16} = 5.55$, $P = 4.40 \times 10^{-5}$; see the supplementary materials for additional details and classification in extrahippocampal ROIs).

Turning to our first central question, we used the "future" classifier to characterize patterns during planning that carry information about future goal locations. Distributed hippocampal activity patterns during planning carried information that significantly distinguished future goal states (classifier accuracy, 29.4%; $t_{16} = 7.54$, $P = 1.19 \times 10^{-6}$) (Fig. 2A). By using neural activity measured during the planning period to classify future goal states, our principal analyses controlled for the contribution of unwanted perceptual and cognitive factors. Specifically, the classification analyses of the planning period targeted representational information that was separated in time from the perception of any past or present goal locations. Consistent with the finding that, in rodents, prospective hippocampal coding for a given location involves reinstatement of the same neural patterns that are present during experience at that location (*17*), a follow-up analysis provided evidence that reinstatement of neural patterns associated with goal arrival occurs during, and contributes to, goal coding during navigational planning (this and other supporting analyses are described in the supplementary materials).

[1]Department of Psychology, Stanford University, Stanford, CA, USA. [2]Department of Psychology, San Jose State University, San Jose, CA, USA. [3]Department of Psychology, New York University, New York, NY, USA. [4]Department of Psychology, University of California–Berkeley, Berkeley, CA, USA. [5]Department of Communication, Stanford University, Stanford, CA, USA. [6]Neurosciences Program, Stanford University, Stanford, CA, USA.
***Corresponding author. Email: thackery@stanford.edu (T.I.B.); awagner@stanford.edu (A.D.W.)**

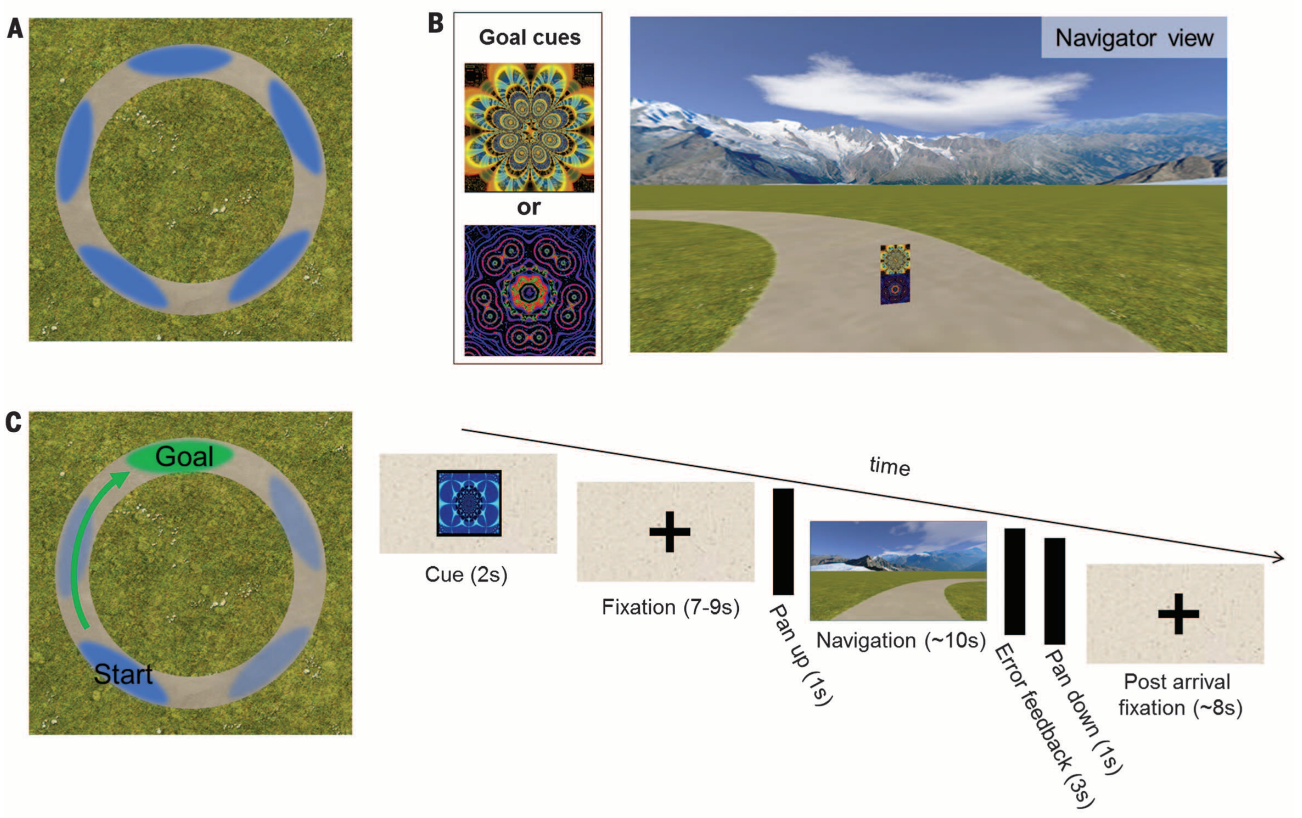


Fig. 1. Task design. (**A**) Overhead view of goal locations (illustrated by blue ellipses) in the virtual environment. (**B**) Example pair of fractals (left) and how fractals appeared at goal locations during day 1 training (right). Fractals were not visible at the locations during day 2 testing. (**C**) Test trial structure. Participants began at one familiar location (blue ellipse), were presented a goal fractal as a cue, and then planned (cue plus fixation periods) and executed navigation to the goal (green ellipse).

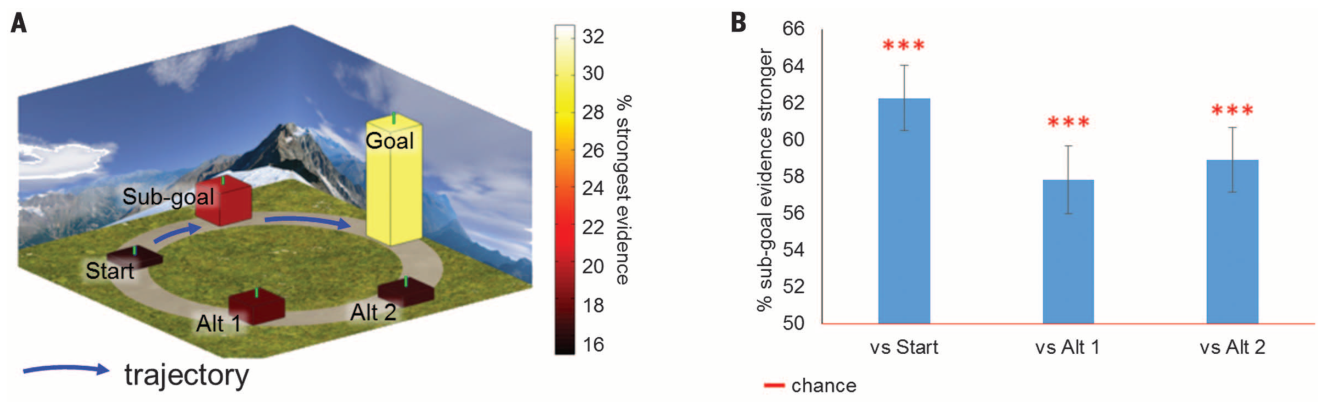


Fig. 2. Hippocampal classifier evidence favors goal and sub-goal (intervening) locations over alternative locations. (**A**) "Future" classifier confusability during planning. Second to the true goal, the classifier most frequently guessed the sub-goal along the planned route (blue arrow). (**B**) Pairwise comparison of sub-goal versus alternative route evidence. Across trials, mean classifier evidence favors the sub-goal over the alternative locations. Error bars reflect the group SEM. ****P* < 0.001.

A second central question is whether the human hippocampus not only supports prospective representation of goal states but also mediates route retrieval during planning. To the extent that planning navigational events incorporates replay of important locations along the route, classifier evidence should favor intervening sub-goals over other nongoal locations. Consistent with this prediction, during navigation planning, the location that was most confusable with the goal was the intervening sub-goal along the optimal route (Fig. 2A and supplementary materials). Direct comparisons of confusability of the goal with the sub-goal versus with the other nongoal locations revealed that the sub-goal was the most favored class (Fig. 2B and supplementary materials).

We also tested whether hippocampal prospective coding is accompanied by future goal–state evidence within a broader cortical network that is thought to subserve the representation and imagery of spatial context features. Specifically, the perirhinal cortex (PRC) may code for item content (environmental cue information) of goal locations (*36*), and the parahippocampal cortex (PHC) and RSC may support planning and future event simulation (*1*) through their putative roles in contextual reinstatement and location coding

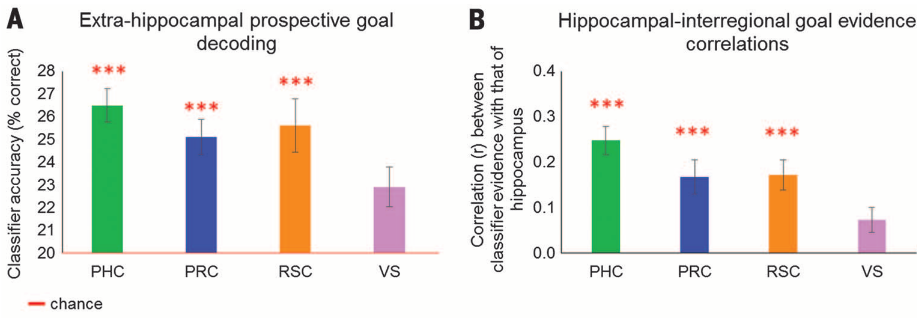


Fig. 3. Prospective evidence in extrahippocampal ROIs. (**A**) Future goal decoding during prospective planning. (**B**) Correlation (Pearson's *r*) between trial-by-trial evidence strength from the "future" classifier in the hippocampus and in extrahippocampal ROIs. Error bars reflect the group SEM. ***$P < 0.001$.

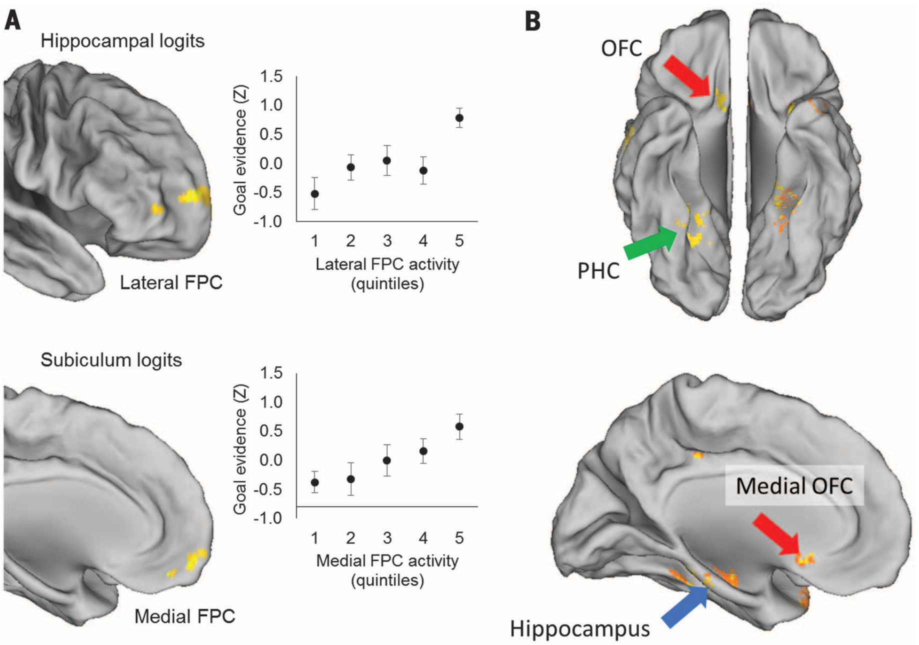


Fig. 4. Prefrontal cortical regions implicated in navigational planning. (**A**) The strength of prospective goal representation in the hippocampus (top) and subiculum (bottom) correlated with univariate activity in the FPC. Plots illustrate the underlying relationship between "future" classifier (goal) evidence (*Z*-score, logits) and the strength of FPC activity extracted from peak voxels. Error bars reflect the group SEM. (**B**) A whole-brain searchlight revealed goal decoding in a core network including the hippocampus, MTL cortex, and OFC. $P < 0.01$, voxel-wise threshold; cluster-corrected $P < 0.05$.

(*8*, *10*, *11*, *33*, *37*). Classification of planning period activity on the basis of the future goal was significantly above chance in each of these regions (Fig. 3A and supplementary materials). VS, which has been implicated in coding motivational signals in space (*19*), exhibited only marginally significant coding for future goal states. Among these a priori ROIs, a whole-brain searchlight revealed local patches in the hippocampus and PHC that exhibited significant goal coding (supplementary materials). Within our PHC, PRC, and RSC ROIs, trial-by-trial classifier evidence for the goal location positively correlated with that in the hippocampus (Fig. 3B and supplementary materials), supporting the hypothesis that their combined representational properties contribute to the multifeatural representation of future spatial contexts.

Top-down, controlled access to episode-specific details in the hippocampus is hypothesized to rely on hippocampal interactions with the PFC (*6, 23, 38*). Computations in the PFC may be important for both expressing goal-relevant mnemonic codes in the hippocampus and integrating hippocampal output into strategic planning. We tested this mechanistic framework by measuring functional connectivity between (i) the hippocampus (more broadly) and hippocampal subfields (more specifically) and (ii) PFC planning period univariate activity and "future" classifier evidence. Planning period activity in the lateral and medial frontopolar cortex (FPC), a region posited to enable prospective expression of memory and help integrate hippocampal output into route plans (*22*, *23*), significantly positively correlated with trial-by-trial "future" classifier evidence in the hippocampus and its subiculum subfield (Fig. 4A). Follow-up analysis of these regions revealed only modest "future" classification in the lateral FPC (that did not survive correction for multiple comparisons; supplementary materials). Instead, the whole-brain searchlight analysis (Fig. 4B) revealed significant "future" classification in the orbitofrontal cortex (OFC), which, critically, is known to connect to and functionally interact with the hippocampus during memory-guided navigation (*11*, *39*). (Methods and complete lists of significant clusters for these analyses are given in the supplementary materials.) Further supporting the importance of functional interaction between the PFC and hippocampal prospective codes in navigational planning, we observed a positive relationship between FPC and (at a modest level) OFC "future" classifier evidence and hippocampal "future" classifier evidence (supplementary materials). Together, these findings suggest that the OFC is part of a hippocampal network that codes for prospective goals and that the FPC plays a role in modulating hippocampal coding, providing cognitive control machinery through which route plans are formed and prospection is achieved (*22*, *23*).

To plan future behavior, humans and animals must be able to represent goals within an environment, as well as to retrieve potential means of reaching these goals. Our data indicate that the hippocampus, interacting with a functionally linked neocortical network (MTL cortex, RSC, and OFC), provides a mechanism for such mental simulation. In particular, our data encompass several important advances: We demonstrate that the human hippocampus contributes to goal-directed navigation, in part through representing future goal states as well as features of the current location (*32*), and, critically, we provide evidence that such prospective retrieval includes episodic simulation of the intended route. Although it remains to be seen whether similar coding and computations occur in more complex large-scale environments, such as those that humans traverse in daily life (*40*), this work bridges the prospective coding of navigational goals in the human hippocampus with related findings in rodents (*3*, *4*, *6*, *17*). Moreover, models of episodic memory and navigation (*6*, *23*, *38*) emphasize the importance of hippocampal-prefrontal interactions for representing navigational events and route planning. Our results provide evidence for an association between prospective hippocampal representations and putative planning processes in the FPC. More broadly, these findings illuminate the mechanistic role of the hippocampus, along with an extended MTL cortex, orbitofrontal, and retrosplenial network, in memory-guided simulation of future events (*1*, *2*). This network, along with the FPC, links look-ahead–like processes with goal-directed planning, which together enable humans to think prospectively.

REFERENCES AND NOTES

1. D. L. Schacter, D. R. Addis, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 1245–1253 (2009).
2. R. L. Buckner, D. C. Carroll, *Trends Cogn. Sci.* **11**, 49–57 (2007).
3. E. R. Wood, P. A. Dudchenko, R. J. Robitsek, H. Eichenbaum, *Neuron* **27**, 623–633 (2000).
4. I. Lee, A. L. Griffin, E. A. Zilli, H. Eichenbaum, M. E. Hasselmo, *Neuron* **51**, 639–650 (2006).
5. D. M. Smith, S. J. Mizumori, *J. Neurosci.* **26**, 3154–3163 (2006).
6. H. T. Ito, S.-J. Zhang, M. P. Witter, E. I. Moser, M.-B. Moser, *Nature* **522**, 50–55 (2015).
7. J. Ferbinteanu, M. L. Shapiro, *Neuron* **40**, 1227–1239 (2003).
8. T. I. Brown, M. E. Hasselmo, C. E. Stern, *Hippocampus* **24**, 819–839 (2014).
9. T. I. Brown, C. E. Stern, *Cereb. Cortex* **24**, 1906–1922 (2014).
10. T. I. Brown, R. S. Ross, J. B. Keller, M. E. Hasselmo, C. E. Stern, *J. Neurosci.* **30**, 7414–7422 (2010).
11. T. I. Brown, R. S. Ross, S. M. Tobyne, C. E. Stern, *Neuroimage* **60**, 1316–1330 (2012).
12. T. I. Brown, A. S. Whiteman, I. Aselcioglu, C. E. Stern, *J. Neurosci.* **34**, 2314–2320 (2014).
13. M. E. Hasselmo, H. Eichenbaum, *Neural Netw.* **18**, 1172–1190 (2005).
14. M. E. Hasselmo, *Neurobiol. Learn. Mem.* **92**, 559–573 (2009).
15. M. E. Hasselmo, C. E. Stern, *Neuroimage* **85**, 656–666 (2014).
16. A. Johnson, A. D. Redish, *J. Neurosci.* **27**, 12176–12189 (2007).
17. A. M. Wikenheiser, A. D. Redish, *Nat. Neurosci.* **18**, 289–294 (2015).
18. D. R. Addis, A. T. Wong, D. L. Schacter, *Neuropsychologia* **45**, 1363–1377 (2007).
19. M. A. A. van der Meer, A. Johnson, N. C. Schmitzer-Torbert, A. D. Redish, *Neuron* **67**, 25–32 (2010).
20. P. Byrne, S. Becker, N. Burgess, *Psychol. Rev.* **114**, 340–375 (2007).
21. N. Burgess, S. Becker, J. A. King, J. O'Keefe, *Philos. Trans. R. Soc. London B Biol. Sci.* **356**, 1493–1503 (2001).
22. N. Ramnani, A. M. Owen, *Nat. Rev. Neurosci.* **5**, 184–194 (2004).
23. H. J. Spiers, S. J. Gilbert, *Front. Hum. Neurosci.* **9**, 125 (2015).
24. J. F. Miller *et al.*, *Science* **342**, 1111–1114 (2013).
25. A. D. Ekstrom *et al.*, *Nature* **425**, 184–188 (2003).
26. L. R. Howard *et al.*, *Curr. Biol.* **24**, 1331–1340 (2014).
27. K. R. Sherrill *et al.*, *J. Neurosci.* **33**, 19304–19313 (2013).
28. H. J. Spiers, E. A. Maguire, *Hippocampus* **17**, 618–626 (2007).
29. C. F. Doeller, C. Barry, N. Burgess, *Nature* **463**, 657–661 (2010).
30. M. J. Chadwick, A. E. J. Jolly, D. P. Amos, D. Hassabis, H. J. Spiers, *Curr. Biol.* **25**, 87–92 (2015).
31. V. Sulpizio, G. Committeri, G. Galati, *Front. Hum. Neurosci.* **8**, 716 (2014).
32. D. Hassabis *et al.*, *Curr. Biol.* **19**, 546–554 (2009).
33. S. A. Marchette, L. K. Vass, J. Ryan, R. A. Epstein, *Nat. Neurosci.* **17**, 1598–1606 (2014).
34. S. A. Marchette, L. K. Vass, J. Ryan, R. A. Epstein, *J. Neurosci.* **35**, 14896–14908 (2015).
35. S. McKenzie, N. T. M. Robinson, L. Herrera, J. C. Churchill, H. Eichenbaum, *J. Neurosci.* **33**, 10243–10256 (2013).
36. J. C. Liang, A. D. Wagner, A. R. Preston, *Cereb. Cortex* **23**, 80–96 (2013).
37. E. M. Aminoff, K. Kveraga, M. Bar, *Trends Cogn. Sci.* **17**, 379–390 (2013).
38. R. Navawongse, H. Eichenbaum, *J. Neurosci.* **33**, 1002–1013 (2013).
39. H. Barbas, G. J. Blatt, *Hippocampus* **5**, 511–533 (1995).
40. T. Wolbers, J. M. Wiener, *Front. Hum. Neurosci.* **8**, 571 (2014).

ACKNOWLEDGMENTS

This work was supported by the National Institute of Mental Health (grant R01-MH076932) and the Wallenberg Network Initiative on Culture, Brain, and Learning. It was also made possible through the support of a grant from the John Templeton Foundation, "Prospective Psychology Stage 2: A Research Competition," to M. Seligman. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. This work was additionally supported by fellowships from the NSF Graduate Research Fellowship Program, the NSF Integrative Graduate Education and Research Traineeship program, and the Natural Sciences and Engineering Research Council of Canada. Raw data are archived on the Stanford Neuroscience Institute server and will be made available upon request.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1323/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
References (*41–58*)

14 December 2015; accepted 12 May 2016
10.1126/science.aaf0784

NEURODEVELOPMENT

Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system

Sueli Marques,[1*] Amit Zeisel,[1*] Simone Codeluppi,[1,2] David van Bruggen,[1] Ana Mendanha Falcão,[1] Lin Xiao,[3,4] Huiliang Li,[3] Martin Häring,[1] Hannah Hochgerner,[1] Roman A. Romanov,[1,5] Daniel Gyllborg,[1] Ana B. Muñoz-Manchado,[1] Gioele La Manno,[1] Peter Lönnerberg,[1] Elisa M. Floriddia,[1] Fatemah Rezayee,[1] Patrik Ernfors,[1] Ernest Arenas,[1] Jens Hjerling-Leffler,[1] Tibor Harkany,[1,5] William D. Richardson,[3] Sten Linnarsson,[1†] Gonçalo Castelo-Branco[1†]

Oligodendrocytes have been considered as a functionally homogeneous population in the central nervous system (CNS). We performed single-cell RNA sequencing on 5072 cells of the oligodendrocyte lineage from 10 regions of the mouse juvenile and adult CNS. Thirteen distinct populations were identified, 12 of which represent a continuum from *Pdgfra*$^+$ oligodendrocyte precursor cells (OPCs) to distinct mature oligodendrocytes. Initial stages of differentiation were similar across the juvenile CNS, whereas subsets of mature oligodendrocytes were enriched in specific regions in the adult brain. Newly formed oligodendrocytes were detected in the adult CNS and were responsive to complex motor learning. A second *Pdgfra*$^+$ population, distinct from OPCs, was found along vessels. Our study reveals the dynamics of oligodendrocyte differentiation and maturation, uncoupling them at a transcriptional level and highlighting oligodendrocyte heterogeneity in the CNS.

Oligodendrocytes ensheath axons in the central nervous system (CNS), allowing rapid saltatory conduction and providing metabolic support to neurons. Although a largely homogeneous oligodendrocyte population is thought to execute these functions throughout the CNS (*1*), these cells were originally described as morphologically heterogeneous (*2*). It is thus unclear whether oligodendrocytes become morphologically diversified during maturation through interactions within the local environment or whether there is intrinsic functional heterogeneity (*3–5*). We analyzed 5072 transcriptomes of single cells expressing markers from the oligodendrocyte lineage, isolated from 10 distinct regions of the anterior-posterior and dorsal-ventral axis of the mouse juvenile and adult CNS (Fig. 1, A and B). Biclustering analysis (*6*) (figs. S1B and S15), hierarchical clustering (Fig. 1C), and differential expression analysis (tables S1 and S2) led to the identification of 13 distinct cell populations. *t*-Distributed stochastic neighbor embedding (t-SNE) (Fig. 2A) supported by pseudotime analysis (fig. S2, A and B) indicated a narrow differentiation path connecting oligodendrocyte precursor cells (OPCs) and myelin-forming oligodendrocytes, which then diversify into six mature states.

Oligodendrocyte precursor cells coexpressed *Pdgfra* and *Cspg4* (Fig. 2B and figs. S1B and S10), and 10% coexpressed cell cycle genes (fig. S2, E and F), consistent with a cell division turnover of 19 days in the juvenile cortex (*7*). Several genes (such as *Fabp7* and *Tmem100*) identified in OPCs were previously associated with astrocytes and radial glia (*6*) (figs. S1B, S3, and S10), consistent with the origin of OPCs from radial glia–like cells, as well as their capacity to generate astrocytes in injury paradigms (*8*).

[1]Laboratory of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-17177 Stockholm, Sweden. [2]Department of Physiology and Pharmacology, Karolinska Institutet, SE-17177 Stockholm, Sweden. [3]Wolfson Institute for Biomedical Research, University College London, Gower Street, London WC1E 6BT, UK. [4]Institute of Neuroscience, Second Military Medical University, 800 Xiangyin Road, Shanghai 200433, China. [5]Department of Molecular Neurosciences, Center for Brain Research, Medical University of Vienna, Vienna, Austria.
*These authors contributed equally to this work. **†Corresponding author. Email: sten.linnarsson@ki.se (S.L.); goncalo.castelo-branco@ki.se (G.C.-B.)**

Differentiation-committed oligodendrocyte precursors (COPs) were distinct from OPCs (because they lacked *Pdgfra* and *Cspg4*) and expressed *Neu4* and genes involved in keeping oligodendrocytes undifferentiated (*Sox6*, *Bmp4*, and *Gpr17*) (*9–11*) (Fig. 2B and figs. S1B, S4, and S10). COPs presented lower levels of cell cycle markers (fig. S2, E and F) while expressing genes involved in migration (*Tns3* and *Fyn*) (fig. S10). Newly formed oligodendrocytes (NFOL1 and NFOL2) expressed genes induced at early stages of differentiation (*Tcf7l2* and *Casr*) (fig. S10) (*12–14*). Whereas *Gpr17* expression decreased in these cells, we observed a peak in levels of *Tcf7l2*, which is involved in oligodendrocyte differentiation (fig. S10) (*15*).

Myelin-forming oligodendrocytes (MFOL1 and MFOL2) expressed genes responsible for myelin formation (*Mal*, *Mog*, *Plp1*, *Opalin*, and *Serinc5*) (fig. S1, A and B). Single-molecule fluorescence RNA in situ hybridization (smFISH) showed that myelin-forming populations (*Ctps*$^+$) were distinct from mature oligodendrocytes (*Klk6*$^+$) (Fig. 2C and fig. S4D). Mature oligodendrocytes (MOL1 to MOL6) expressed late oligodendrocyte differentiation genes (*Klk6* and *Apod*) (*12*), as well as genes present in myelinating cells (*Trf* and *Pmp22*) (fig. S1B).

We identified a second *Pdgfra*$^+$ population—vascular and leptomeningeal cells (VLMCs)—distinct from OPCs and segregated from all oligodendrocyte lineage cells (Figs. 1C and 2A). This population was also found when sorting green fluorescent protein (GFP)–positive cells from *Pdgfra*–histone 2B (H2B)–GFP (*16*) and *Pdgfra*-Cre-RCE (LoxP-GFP) mice (*17*) (fig. S2C). These cells exhibited low levels of *Cspg4* (NG2) (Fig. 2B) and specifically expressed *Lum* (Fig. 2B and fig. S4), markers of the pericyte lineage (*Vtn* and *Tbx18*) (Fig. 2B and figs. S1B and S2D), and laminins and collagens characteristic of the basal lamina. *Pdgfra*$^+$ and *Sox10*$^-$ VLMCs were localized on blood vessels (Fig. 2D and figs. S4 and S11, A and B) and meninges (fig. S11, A to C). In contrast, COL1A1$^-$ and PDGFRA$^+$ OPCs were distributed in the parenchyma, in close association but not overlapping with the vasculature (Fig. 2D and fig. S11B) (*18*). VLMCs specifically exhibited markers present in transcriptomes of OPCs isolated based on PDGFRA$^+$ immunoreactivity (fig. S3) (*14*), most likely previously assigned to OPCs due to copurification.

We retrieved the 50 genes that better differentiate every branch of the dendrogram plot (Fig. 1C) and investigated their putative function by gene ontology analysis (figs. S6 to S9 and tables S1 and S2). COPs were enriched in cell fate commitment and adhesion genes, whereas newly formed oligodendrocytes (NFOL1 and NFOL2) already presented genes involved in steroid biosynthesis, ensheathment of neurons, and cell projection organization (fig. S7). These populations exhibited distinct expression of *Tcf7l2*, *Itpr2*, *Tmem2*, and *Pdgfa* (Fig. 3A and fig. S4). ITPR2, encoding an intracellular Ca^{2+} channel, was more specific to oligodendrocytes than TCF7L2 and exhibited close to 100% overlap with SOX10-positive cells (fig. S5, A and D). We observed that

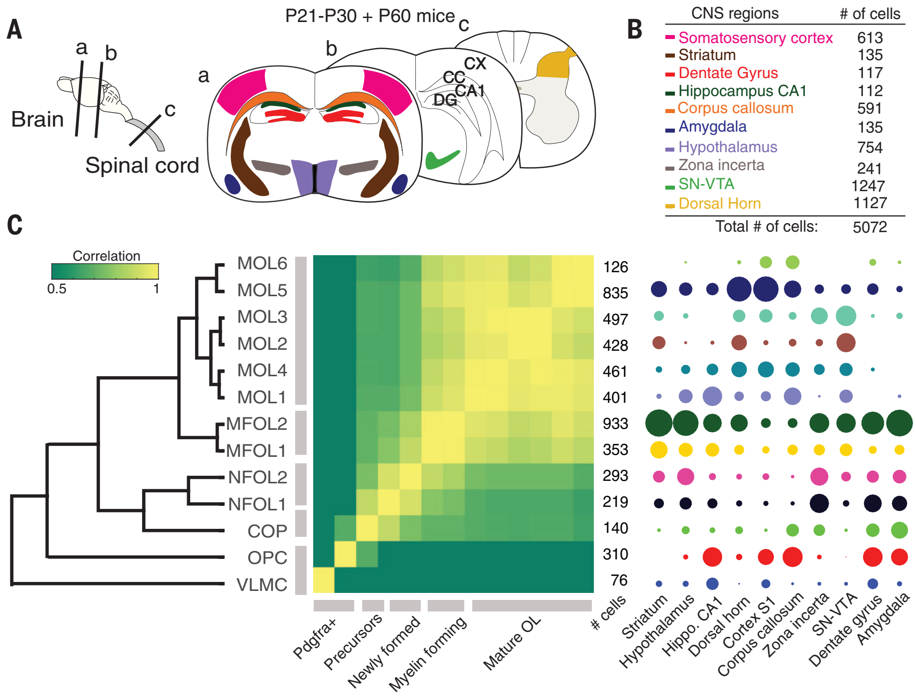


Fig. 1. Single-cell RNA sequencing analysis of 5072 cells expressing markers of the oligodendrocyte lineage in 10 regions of the mouse CNS. (**A**) Targeted regions. CX, cortex; CC, corpus callosum; CA1, CA1 hippocampus; DG, dentate gyrus. (**B**) Number of cells analyzed for each region. SN-VTA, substantia nigra ventral tegmental area. (**C**) Hierarchical clustering (left), correlation matrix (middle), and subclass abundances by region (right). OL, oligodendrocytes.

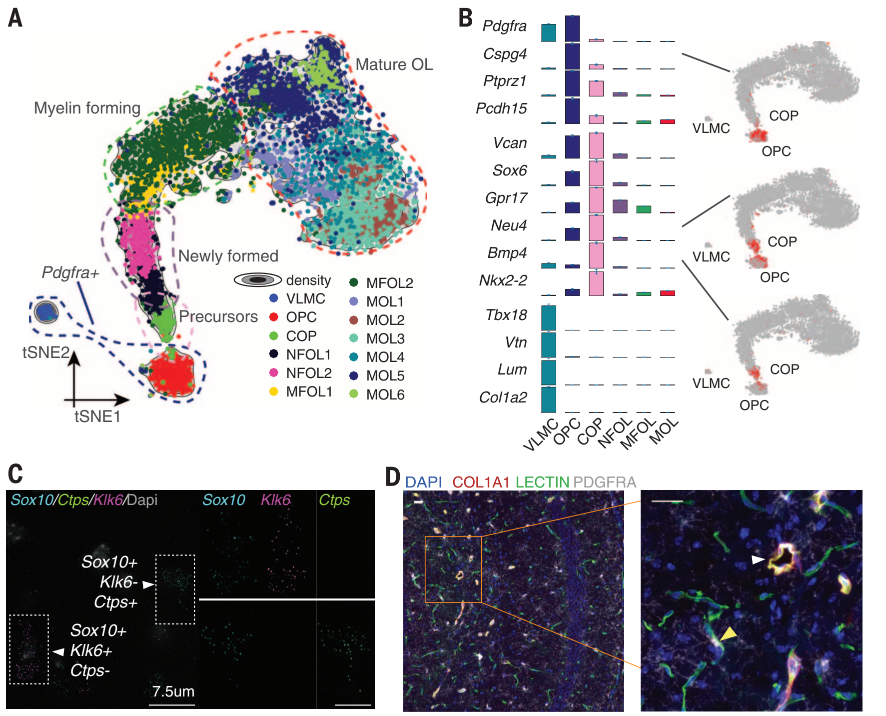


Fig. 2. Oligodendrocyte cell states in the continuous maturation process from precursors to mature cells. (**A**) *t*-Distributed stochastic neighbor embedding projection showing the trajectory from OPCs to mature oligodendrocytes. (**B**) Average (±SEM) expression of marker genes for OPCs, COPs, and VLMCs. Representative markers are overlaid on the t-SNE map (gray, low expression; red, high expression). (**C**) Results of smFISH for *Sox10*, *Ctps* (MFOL marker), and *Klk6* (MOL marker) confirm that these populations are distinct. DAPI, 4′,6-diamidino-2-phenylindole. Scale bar, 7.5 μm. (**D**) Immunohistochemistry of COL1A1 (VLMCs), PDGFRA (OPCs and VLMCs), and Tomato LECTIN (blood vessels) in the brain at P21. White arrowhead, VLMCs; yellow arrowhead, OPCs (COL1A1$^-$). Scale bars, 25 μm.

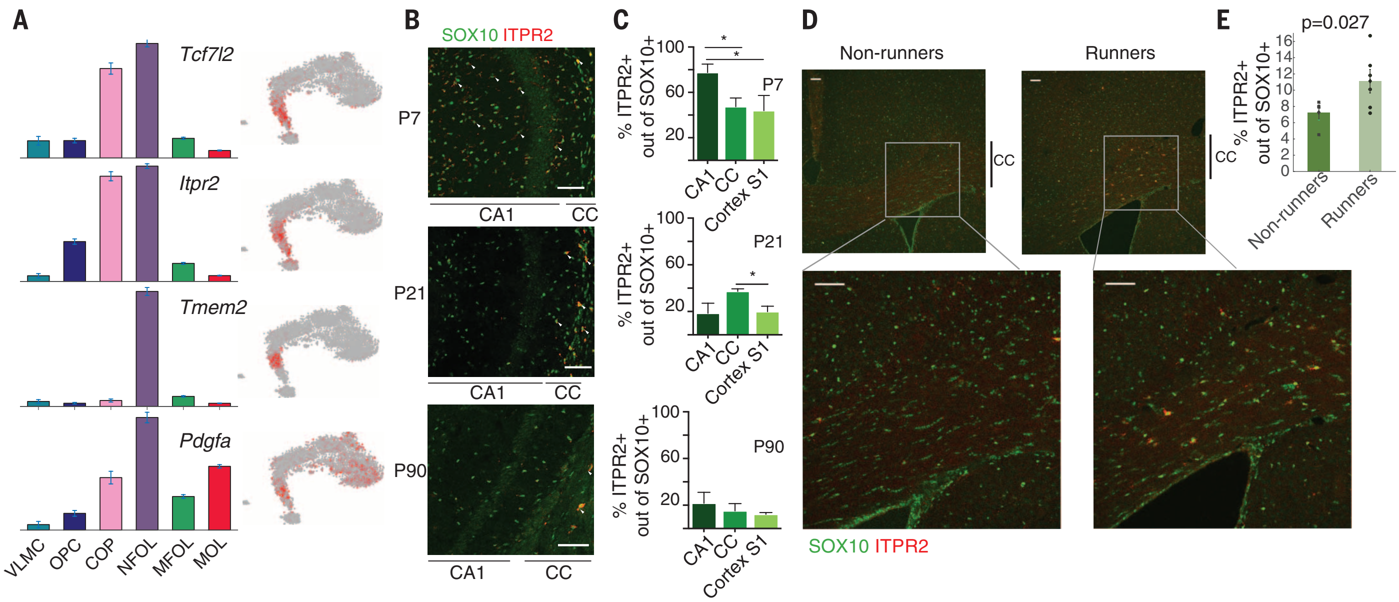


Fig. 3. ITPR2$^+$ oligodendrocytes are present in regions of active differentiation and increase in mice undergoing learning in the complex wheel paradigm. (**A**) Average (±SEM) expression level of *Tcf7l2*, *Itpr2*, *Tmem2*, and *Pdgfa* along the oligodendrocyte lineage. (**B** and **C**) Immunohistochemistry and quantification of ITPR2$^+$ out of SOX10$^+$ cells in the brain at P7, P21, and P90. One-way analysis of variance with Tukey's multiple comparison test (*$P < 0.05$, $n = 3$ mice per time point). Error bars indicate SEM. (**D** and **E**) Immunohistochemistry and quantification of ITPR2$^+$ out of SOX10$^+$ cells in the corpus callosum of P60 nonrunners (squares) versus runners (circles) after 2 days in the complex wheel-learning paradigm (one-tailed Student's t test). Scale bars in (D), 75 μm. Error bars in (E) indicate SEM.

ITPR2 immunoreactive cells were distinct from PDGFRA$^+$ OPCs (fig. S5B), and lineage tracing confirmed that ITPR2$^+$ cells are the progeny of OPCs (figs. S2C and S5C). Of the OPC-derived *Pdgfra*-H2B-GFP$^+$ cells, 22 ± 2 and 25 ± 1.5% were positive for ITPR2 in the somatosensory (S1) cortex and the CA1 hippocampus, respectively, at postnatal day 21 (P21), whereas 43 ± 3.7% double-positive cells were found in the corpus callosum (fig. S5C). The percentage of ITPR2$^+$, Sox10$^+$ cells in the corpus callosum remained within the same range at P7 (47 ± 4%) and P21 (37 ± 1%) (Fig. 3C). Of the SOX10$^+$ oligodendrocytes, 77 ± 4 and 48 ± 7% were positive for ITPR2 at P7 in the CA1 hippocampus and the S1 cortex, respectively, and decreased to less than 20% thereafter (Fig. 3, B and C). This distribution of ITPR2$^+$ oligodendrocytes correlates with active and prolonged differentiation in the juvenile rat corpus callosum (*19*). These tissues still maintained 10 to 20% ITPR2$^+$ cells at adult stages (P90 in Fig. 3C).

To investigate the potential function of ITPR2$^+$ cells in the adult brain, we analyzed their dynamics in the corpus callosum of mice engaged in motor learning on the complex wheel, a process that requires active myelination (*20*). In this paradigm, running on the wheel leads to an increase in the number of proliferating OPCs after 4 days, followed by an increase in oligodendrocytes after 8 days (*20*). However, increased motor skills were already apparent after 2 days in wild-type mice, but not in mutant mice that were unable to synthesize new myelin (*20*), suggesting that oligodendrocyte lineage cells already contribute to learning within the first 2 days. We found that the number of ITPR2$^+$, SOX10$^+$ cells increased by ~50% in mice that ran on the complex wheel for 2 days, as compared with nonrunners (Fig. 3, D and E). Thus, novel motor activity might trigger rapid differentiation of OPCs into ITPR2$^+$ committed precursors and newly formed oligodendrocytes that contribute to early learning by facilitating electrical transmission, either through the initiation of myelination or some other premyelinating function.

We were unable to identify region- or age (juvenile versus adult)–specific subpopulations of OPCs in our data set (Figs. 2A and 4, A and B). Nevertheless, 16% of the juvenile OPCs were in the cell cycle [as determined by the simultaneous expression of more than two cell cycle markers (fig. S2F)], compared with ~3% of the adult OPCs. Similarly, COPs and newly formed oligodendrocytes were present in all regions in juvenile mice (Figs. 1C and 4A), revealing a common trajectory of differentiation between the various regions (Fig. 2A). These populations were also observed in the adult corpus callosum and the S1 cortex, albeit in considerably lower numbers compared with those seen in juvenile mice (Fig. 4B). On the basis of the distribution of cell types in the juvenile mice, we classified regions as immature (anterior regions such as the amygdala and hippocampus), intermediate (corpus callosum, zona incerta, striatum, and hypothalamus), and mature (cortex and posterior regions such as the dorsal horn and the substantia nigra ventral tegmental area) (Fig. 4A and fig. S12). These regional variations could result from different timing of oligodendrocyte maturation during postnatal development (*21*, *22*). Myelination first starts in the rat in posterior regions (dorsal horn) around P7, whereas in anterior regions of the CNS (amygdala, hippocampus, striatum, and cortex) it occurs between P21 and P28 (*23*).

Different regions of the CNS were populated by diverse mature oligodendrocytes (Fig. 1C and fig. S12). Although some populations, such as MOL5, were present throughout the regions, other mature oligodendrocytes were enriched in certain regions (fig. S12). Some of these mature oligodendrocyte populations might be intermediate stages or have specific functions in juvenile mice but then disappear in adulthood. Subsets of MOL5 and MOL6 were mainly present in the S1 cortex and corpus callosum in the adult mice (Fig. 4B). Because MOL5 was already present in several regions of the juvenile CNS (Fig. 1C and fig. S12), final maturation of oligodendrocytes might already be achieved in the juvenile mice in certain regions, such as the dorsal horn, but only in adulthood in others, such as the corpus callosum.

Gene ontology analysis indicated a divergence already at the stage of myelin formation (fig. S8 and tables S1 and S2). Although mature oligodendrocyte populations shared the expression of many genes, some were differentially enriched within populations (fig. S8 and tables S1 and S2), indicating segregation of MOL1 to MOL4, enriched in lipid biosynthesis and myelination genes (*Far1* and *Pmp22*) from MOL5 and MOL6 (adult), enriched for synapse parts such as *Grm3* (metabotropic glutamate receptor, mainly enriched in MOL6) and *Jph4*. We confirmed the presence of GRM3 in the oligodendrocyte lineage (*Pdgfra*-Cre-RCE) and specifically in CC1$^+$ mature oligodendrocytes in the juvenile cortex (fig. S11D). Even within MOL1 to MOL4, which were enriched

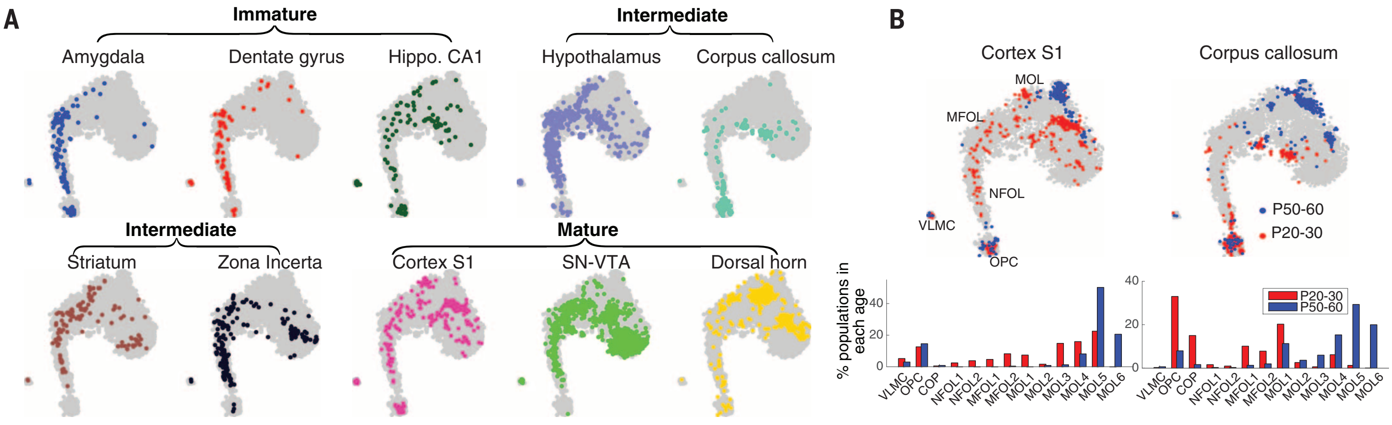


Fig. 4. Region- and age-specific distribution of mature oligodendrocytes. (**A**) *t*-Distributed stochastic neighbor embedding projections, as in Fig. 2A, with colored dots representing cells from each of the 10 CNS regions analyzed. (**B**) Age distribution of oligodendrocyte populations in the S1 cortex and corpus callosum. Bar plots show the percentage of each population by age. Red, juvenile brain; blue, adult brain.

in myelin-related genes, specific populations (such as MOL3) are more likely to be involved in synaptic activity (fig. S9 and tables S1 and S2). Optic nerve oligodendrocytes can form axon-myelinic synapses, responding to axonal action potentials via glutamate ionotropic *N*-methyl-D-aspartate receptors (*24*). We analyzed the expression of ionotropic and metabotropic glutamate receptors and other ion channels, including transient receptor potential (TRP) (*25*) and potassium channels (fig. S14). Although most glutamate receptor subunits were expressed throughout oligodendrocyte lineage cells, there was preferential expression in some populations, with single cells displaying combinations of subunits that might determine function. Potassium channels and TRPs were also expressed in a cell type–specific manner, displaying a scattered distribution within populations (fig. S14). Thus, the communication of mature oligodendrocytes with neighboring neurons might be mediated through specific receptors and channels, following synaptic input or vesicular release.

Our study provides a high-resolution view of the transcriptional landscape of a single neural subtype across multiple regions of the CNS and indicates a transcriptional continuum between oligodendrocyte populations, with a subset representing distinct but nevertheless connected stages in the maturation path from OPCs to mature oligodendrocytes (fig. S16). Initial differentiation was uniform throughout the CNS, whereas mature oligodendrocyte subtype specification occurred later at postnatal stages and in a region-specific manner. Each brain region appears to optimize its circuitry by representation of distinct proportions and combinations of mature oligodendrocytes. Our data also indicate that ITPR2$^+$ oligodendrocytes are involved in rapid myelination in complex motor learning and thus might be relevant in other active maturation and myelination processes, such as remyelination in disease or lesion paradigms. Nonproliferative *Nkx2.2*$^+$ precursors that have a profile consistent with these cells (fig. S10) have been observed in lesions of patients with multiple sclerosis (*26*). Therefore, by establishing oligodendrocytes as a transcriptionally heterogeneous cell lineage, our study might lead to new insights into the etiology of myelin disorders, such as multiple sclerosis, and might suggest novel targets for their treatment.

REFERENCES AND NOTES

1. N. Kessaris *et al.*, *Nat. Neurosci.* **9**, 173–179 (2006).
2. P. d. Rio-Hortega, *Mem. Real Soc. Esp. Hist. Nat.* **14**, 5–122 (1928).
3. G. S. Tomassy *et al.*, *Science* **344**, 319–324 (2014).
4. R. B. Tripathi *et al.*, *J. Neurosci.* **31**, 6809–6819 (2011).
5. M. E. Bechler, L. Byrne, C. Ffrench-Constant, *Curr. Biol.* **25**, 2411–2416 (2015).
6. A. Zeisel *et al.*, *Science* **347**, 1138–1142 (2015).
7. K. M. Young *et al.*, *Neuron* **77**, 873–885 (2013).
8. L. Dimou, M. Götz, *Physiol. Rev.* **94**, 709–737 (2014).
9. C. C. Stolt *et al.*, *Dev. Cell* **11**, 697–709 (2006).
10. J. Samanta, J. A. Kessler, *Development* **131**, 4131–4142 (2004).
11. Y. Chen *et al.*, *Nat. Neurosci.* **12**, 1398–1406 (2009).
12. J. D. Cahoy *et al.*, *J. Neurosci.* **28**, 264–278 (2008).
13. V. A. Swiss *et al.*, *PLOS ONE* **6**, e18088 (2011).
14. Y. Zhang *et al.*, *J. Neurosci.* **34**, 11929–11947 (2014).
15. F. Ye *et al.*, *Nat. Neurosci.* **12**, 829–838 (2009).
16. R. A. Klinghoffer, T. G. Hamilton, R. Hoch, P. Soriano, *Dev. Cell* **2**, 103–113 (2002).
17. K. Roesch *et al.*, *J. Comp. Neurol.* **509**, 225–238 (2008).
18. H. H. Tsai *et al.*, *Science* **351**, 379–384 (2016).
19. K. Hamano *et al.*, *Brain Res. Dev. Brain Res.* **108**, 287–293 (1998).
20. I. A. McKenzie *et al.*, *Science* **346**, 318–322 (2014).
21. H. C. Kinney, B. A. Brody, A. S. Kloman, F. H. Gilles, *J. Neuropathol. Exp. Neurol.* **47**, 217–234 (1988).
22. B. A. Brody, H. C. Kinney, A. S. Kloman, F. H. Gilles, *J. Neuropathol. Exp. Neurol.* **46**, 283–301 (1987).
23. J. C. Coffey, K. W. McDermott, *J. Neurocytol.* **26**, 149–161 (1997).
24. I. Micu *et al.*, *Exp. Neurol.* **276**, 41–50 (2016).
25. N. B. Hamilton, K. Kolodziejczyk, E. Kougioumtzidou, D. Attwell, *Nature* **529**, 523–527 (2016).
26. T. Kuhlmann *et al.*, *Brain* **131**, 1749–1758 (2008).

ACKNOWLEDGMENTS

We thank P. Soriano (Mount Sinai, New York) for the *Pdgfra*-H2B-GFP mouse; C. Göritz for discussions; and Samudyata, G. Chen, A. Nanni, and J. Söderlund for additional support. P.E. was supported by the Swedish Research Council (VR) (Medicine and Health), the Swedish Cancer Society, and the Wallenberg Scholar and Söderberg Foundation. E.A. was supported by VR (Linnaeus Center in Developmental Biology for Regenerative Medicine, grants 2011-3116 and 2011-3318), the Swedish Foundation for Strategic Research (Successful Research Leaders program), the European Union (EU) (NeuroStemCellRepair and DDPDGENES), and the Karolinska Institutet (KI) (SFO Stem Cells and Regenerative Medicine). J.H.-L. was supported by VR, StratNeuro, Hjärnfonden, and EU FP7/Marie Curie Actions. R.A.R. was supported by the European Molecular Biology Organization (ALTF reference 596-2014) and EU FP7 (Marie Curie Actions, EMBOCOFUND2012, grant GA-2012-600394). T.H. was supported by VR, Hjärnfonden, the Petrus and Augusta Hedlunds Foundation, the Novo Nordisk Foundation, the European Research Council (ERC) (SECRET-CELLS), and the EU (PAINCAGE). W.D.R. was supported by the ERC (grant 293544) and the Wellcome Trust (grants 100269/Z/12/Z and 108726/Z/15/Z). L.X. was supported by the National Natural Science Foundation of China (grant 31471013). H.L. was supported by a New Investigator Award from the UK Biotechnology and Biological Sciences Research Council (BB/L003236/1). A.Z. was supported by the Human Frontier Science Program. S.L. was supported by the ERC (BRAINCELL grant 261063), VR (STARGET), the Wellcome Trust (grant 108726/Z/15/Z), and the EU (FP7/DDPDGENES). A.M.F. was supported by the European Committee for Treatment and Research of Multiple Sclerosis. G.C.-B. was supported by VR (grant 2015-03558), the EU (FP7/Marie Curie Integration Grant, EPIOPC), Hjärnfonden, the Swedish Society of Medicine, Åke Wiberg, Clas Groschinsky, the Petrus och Augusta Hedlunds Foundations, and KI. Data sets were deposited in the Gene Expression Omnibus (accession no. GSE75330) and in a web interface at http://linnarssonlab.org/oligodendrocytes/. The supplementary materials contain additional data. S.M., A.Z., H.L., W.D.R., S.L., and G.C.-B. designed the experiments. P.E., E.A., J.H.-L., T.H., W.D.R., S.L., and G.C.-B. (senior authors) obtained funding. S.M., A.Z., S.C., H.H., R.A.R., D.G., M.H., A.B.M.-M., G.L.M., F.R., H.L., L.X., and E.M.F. performed experiments. L.X., H.L., and W.D.R. have priority of observation of the rapid differentiation of oligodendrocytes in the complex motor wheel paradigm. S.M., A.Z., D.v.B., A.M.F., G.L.M., and P.L. analyzed the data. S.M., A.Z., S.L., and G.C.-B. wrote the paper, with the assistance of and proofreading by all authors.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1326/suppl/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 to S3
References (*27–34*)

9 March 2016; accepted 10 May 2016
10.1126/science.aaf6463

TRANSCRIPTION

Structural basis of transcription activation

Yu Feng, Yu Zhang, Richard H. Ebright*

Class II transcription activators function by binding to a DNA site overlapping a core promoter and stimulating isomerization of an initial RNA polymerase (RNAP)–promoter closed complex into a catalytically competent RNAP-promoter open complex. Here, we report a 4.4 angstrom crystal structure of an intact bacterial class II transcription activation complex. The structure comprises *Thermus thermophilus* transcription activator protein TTHB099 (TAP) [homolog of *Escherichia coli* catabolite activator protein (CAP)], *T. thermophilus* RNAP σ^A holoenzyme, a class II TAP-dependent promoter, and a ribotetranucleotide primer. The structure reveals the interactions between RNAP holoenzyme and DNA responsible for transcription initiation and reveals the interactions between TAP and RNAP holoenzyme responsible for transcription activation. The structure indicates that TAP stimulates isomerization through simple, adhesive, stabilizing protein-protein interactions with RNAP holoenzyme.

Simple bacterial transcription activators—those that stimulate transcription from a single DNA site without other factors—are divided into two classes (*1*–*3*). Class I transcription activators, typified by *Escherichia coli* catabolite activator protein (CAP) at the *lac* promoter, stimulate transcription by binding to a specific DNA site upstream of a core promoter and facilitating binding of RNA polymerase (RNAP) holoenzyme to form an RNAP-promoter closed complex (RPc) (*1*–*3*). Class II transcription activators, typified by *E. coli* CAP at the *gal* promoter, stimulate transcription by binding to a specific DNA site overlapping a core promoter and facilitating conversion of RPc into a catalytically competent RNAP-promoter open complex (RPo) containing ~13 base pairs (bp) of unwound DNA ("transcription bubble") (*1*–*3*). A 20 Å–resolution electron microscopy structure of a class I transcription activation complex has been reported (*4*), but no structure of a class II transcription activation complex previously has been reported. Here, we determine the 4.4 Å–resolution crystal structure of a class II transcription activation complex comprising *Thermus thermophilus* transcription activator protein TTHB099 (TAP) (a thermophilic sequence, structural, and functional homolog of *E. coli* CAP) (*5*); *T. thermophilus* RNAP σ^A holoenzyme; a class II TAP-dependent promoter; and the ribotetranucleotide primer UpCpGpA (TAP-Rpo) (table S1, Fig. 1, and figs. S1 and S2).

To obtain a structure of TAP-RPo, we used a nucleic-acid scaffold corresponding to positions –57 to +15 of a class II TAP-dependent promoter (positions numbered relative to transcription start site) (Fig. 1A and fig. S1). The scaffold contained a consensus DNA site for TAP centered between positions –41 and –42 (same position as DNA site for *E. coli* CAP in *gal* promoter) (*1*, *2*), a near-consensus extended –10 element (*3*), a consensus –10 element (*3*), a consensus discriminator element (*3*), a consensus core recognition element (*3*), a 13-bp transcription bubble (maintained in the unwound state by having noncomplementary sequences on nontemplate and template strands), and UpCpGpA.

In the structure of TAP-RPo, TAP interacts with DNA, RNAP holoenzyme interacts with DNA, and TAP and RNAP holoenzyme make protein-protein interactions (Fig. 1, B and C). The structure of TAP-DNA in TAP-RPo is superimposable on the structure of CAP-DNA (*6*), corroborating that TAP is a homolog of CAP (Fig. 1D). The structure of RPo in TAP-RPo is essentially superimposable on structures of RPo (*7*–*10*) [neglecting RNAP α subunit C-terminal domain (αCTD), which was not resolved in previous structures], indicating that interactions between the class II activator and RPo do not substantially alter the conformation of RPo (Fig. 1E).

RNAP contains two copies of αCTD, each of which is connected to the rest of RNAP through a flexible linker (*1*–*3*). In the structure of TAP-RPo, one αCTD (probably αCTD^{I}) (fig. S3) interacts with TAP, and the other αCTD (probably αCTD^{II}) (fig. S3) makes no interactions (Fig. 1, B and C). In the crystal, the second αCTD is constrained by lattice contacts (i.e., contacts with TAP in an adjacent molecule of TAP-RPo in the lattice) (fig. S4). In solution, this αCTD would be free to adopt other positions.

The structure defines the interactions between RNAP holoenzyme and DNA that mediate promoter recognition and promoter unwinding in transcription initiation (Figs. 1 and 2) and the interactions between TAP and RNAP holoenzyme that mediate transcription activation (Figs. 1, 3, and 4).

TAP and σ conserved region σR4 "corecognize" the promoter –35 region, contacting the same DNA segment from different faces of the DNA helix (Figs. 1, B and C, and 2A). The general mode of interaction of σR4 with –35-region DNA in TAP-RPo—binding of the second α helix of the σR4 helix-turn-helix motif in the DNA major groove—is the same as in RPo (Fig. 2A and fig. S5) (*8*–*10*), but, due to DNA distortion by TAP, –35-region DNA is rotated ~20° away from σR4 (fig. S5). This rotation decreases the number of σR4 residues that contact DNA bases from 3 to 2 and decreases the number of contacted DNA bases from 4 to 2, providing a structural explanation for the observation that, although –35-region DNA sequences are recognized in class II activator-dependent transcription, the recognition specificity is less and the number of recognized bases is smaller than in activator-independent transcription (*11*). Two σR4 residues are positioned to make contacts with DNA bases that potentially enable sequence readout (Fig. 2A). Substitution of these residues reduces RPo formation, verifying their importance (Fig. 2A).

σ conserved region σR3 interacts with the promoter extended –10 region (Figs. 1, B and C, and 2B) (*8*–*10*). Three σR3 residues are positioned to make contacts with DNA bases (Fig. 2B and fig. S6). Substitution of these residues reduces RPo formation, verifying their importance (Fig. 2B).

σ conserved region σR2 interacts with the promoter –10 element at the "upstream fork junction" where DNA unwinding occurs to form the transcription bubble (Figs. 1E and 2C). σR2 interacts with the first position of the –10 element (–12) as double-stranded DNA (dsDNA) and the second through sixth positions of the –10 element (–11 through –7) as nontemplate-strand single-stranded DNA (ssDNA) (Figs. 1E and 2C). σR2 Trp^{433} (numbered as in *E.* coli σ^{70}) is positioned to stack on the nontemplate-strand base of base pair –12, forming a "wedge" that forces the nontemplate-stand –11 base to unstack and flip outside the DNA helix (*9*), where it is captured by binding within a pocket formed by residues of σR2 (Fig. 2C and fig. S6) (*7*–*10*). σR2 Arg^{436} is positioned to stack on the template-strand base of base pair –12, forming an analogous "wedge" that forces the template-strand –11 base to unstack and flip outside the DNA helix, where it is captured within a channel formed by residues of RNAP, σR2, and σR3.2 that leads into the RNAP active-center cleft (Fig. 2C and figs. S6 and S7). Substitution of Trp^{433} or Arg^{436} results in defects in RPo formation, verifying their importance (Fig. 2C). A second pair of residues, Gln^{437} and Thr^{440}, are positioned to make direct contacts with the nontemplate- and template-strand bases of the –12 base pair, providing a structural explanation for the observation that substitution of these residues alters specificity at –12 (Fig. 2C and fig. S6) (*12*).

σ conserved region σR1.2 interacts with nontemplate-strand ssDNA of the discriminator element, as described previously (Fig. 1E) (*7*–*10*). RNAP core interacts with the nontemplate-strand

Waksman Institute and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA.
*Corresponding author. Email: ebright@waksman.rutgers.edu

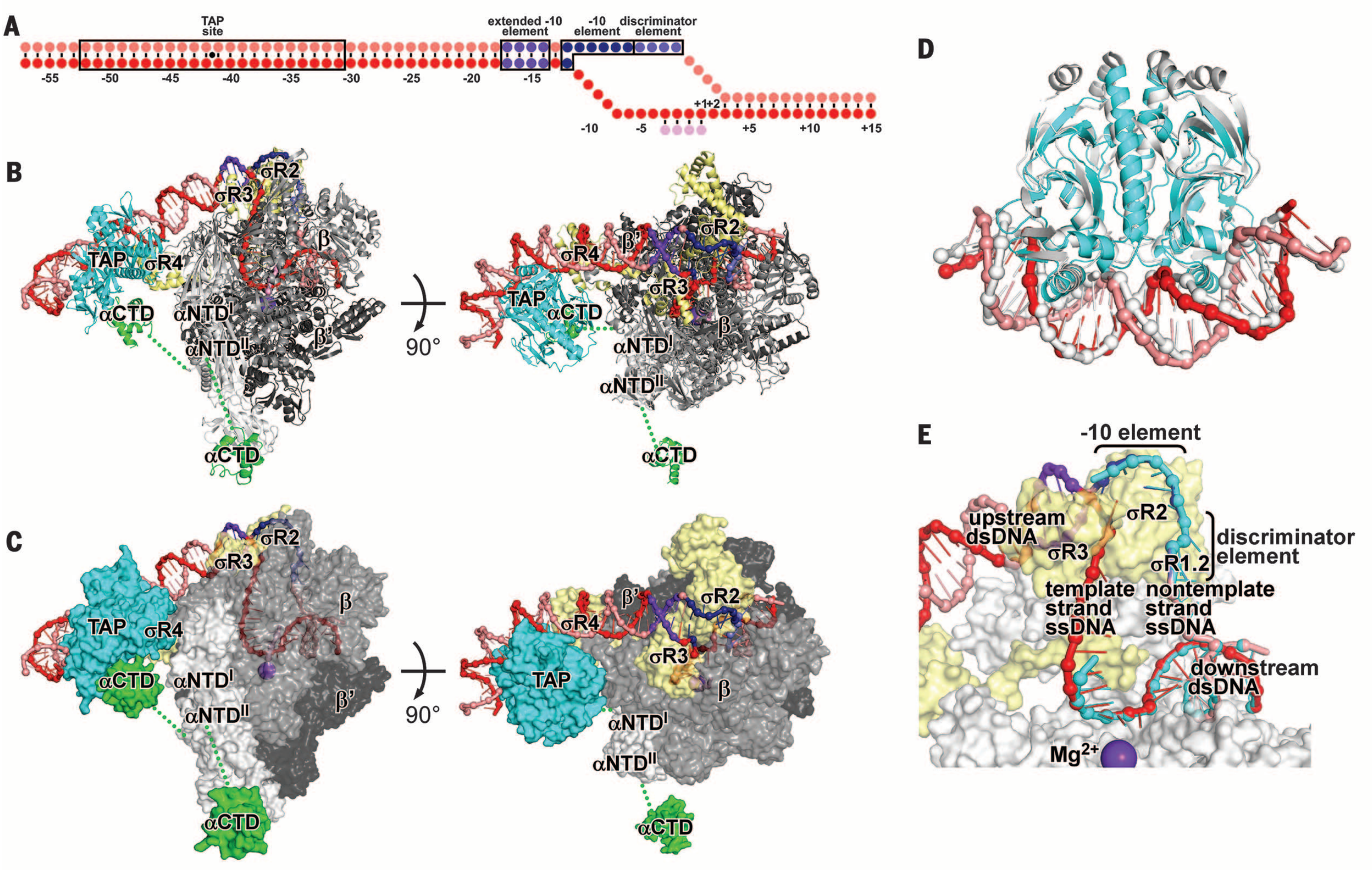


Fig. 1. Structure of TAP-RPo. (**A**) Nucleic-acid scaffold. Pink, nontemplate strand; red, template strand; magenta, UpCpGpA; violet, extended –10 element; blue, –10 element; light blue, discriminator element. (**B** and **C**) TAP-RPo [ribbons in (B); surfaces in (C); β' nonconserved region omitted for clarity]. Cyan, TAP; yellow, σ; white, green, gray, and dark gray, RNAP αNTD, αCTD, β, and β', respectively. Other colors as in (A). Dashed lines, αNTD-αCTD linkers. (**D**) Comparison of TAP-DNA in TAP-RPo [colors as in (B) and (C)] to CAP-DNA (gray) (*6*). (**E**) Comparison of transcription bubble and downstream dsDNA in TAP-RPo [colors as in (B) and (C)] to corresponding DNA segments in RPo (cyan) (*7*).

ssDNA of the core recognition element, template-strand ssDNA of the transcription bubble, and downstream dsDNA, as described previously (Fig. 1E) (*7*–*10*).

Genetic and biochemical experiments indicate that class II transcription activation by *E. coli* CAP involves three sets of protein-protein interactions: (i) activating region 1 (AR1) interacts with αCTD, (ii) activating region 2 (AR2) interacts with a species-specific insertion in RNAP α^{I} subunit N-terminal domain (αNTD^{I}) (162-165 determinant), and (iii) activating region 3 (AR3) interacts with σR4 (*1*, *2*).

In TAP-RPo, a surface of TAP corresponding to AR2 approaches αNTD^{I} and contacts the RNAP β flap (Fig. 3A). Three residues of TAP AR2 are positioned to make direct contacts with three residues of RNAP β subunit (Fig. 3B). TAP Glu^{77} and RNAP β Arg^{735} are positioned to form a salt bridge in the AR2-RNAP interface (Fig. 3B). Charge-reversal substitution of either residue decreases TAP-dependent transcription, and charge-reversal substitution of both residues, which recreates a salt bridge, restores TAP-dependent transcription, confirming the importance of the inferred interaction (Fig. 3B). Homology modeling of CAP-RPo based on TAP-RPo indicates that CAP AR2 is positioned to contact the αNTD^{I} 162-165 determinant (a species-specific insertion present in *E. coli* RNAP but not in *T. thermophilus* RNAP) (fig. S8, A and B), consistent with previous work (*13*). Homology modeling indicates that CAP AR2 also is positioned to contact the RNAP β flap (fig. S8, A and B). Substitution of the inferred interacting residues decreases CAP-dependent transcription, indicating that the inferred interactions occur and are important (fig. S8B).

In TAP-RPo, a surface of TAP corresponding to AR3 contacts σR4 α helices 4 and 5 and the RNAP β flap-tip α helix (Fig. 3A). Eight predominantly negatively charged residues of TAP AR3 are positioned to interact with six predominantly positively charged residues of σR4 and three predominantly positively charged residues of the β flap-tip α helix (Fig. 3C). TAP Glu^{15} is positioned to form a salt bridge with a σR4 Arg residue at the center of the interface (Fig. 3C). Charge-reversal substitution of either residue decreases TAP-dependent transcription, and charge-reversal substitution of both residues, recreating a salt bridge, restores TAP-dependent transcription, indicating that the interactions occur and are important (Fig. 3C). Homology modeling of CAP-RPo based on TAP-RPo predicts equivalent interactions between seven predominantly negatively charged residues of CAP AR3 and five predominantly positively charged residues of σR4 and one residue of the β flap-tip (fig. S8, A and C), consistent with previous work (*14*, *15*).

In TAP-RPo, the surface of TAP corresponding to AR1 makes no interactions, and, instead, a different surface of TAP, here designated "activating region 4" (AR4), interacts with αCTD (Figs. 1, B and C, and 4A). The interface between TAP AR4 and αCTD is large (300 $Å^2$) (Fig. 4B). Nine residues of TAP AR4 are positioned to make direct contacts with eight residues of αCTD (Fig. 4B). Substitution of residues implicated in TAP AR4-αCTD interaction results in defects in TAP-dependent transcription (Fig. 4B). TAP-αCTD interactions differ from CAP-αCTD interactions not only in the identities of the activating regions (AR4 in TAP; AR1 in CAP) but also in the fact that TAP interacts with αCTD not bound to DNA, whereas CAP interacts with αCTD bound to DNA immediately upstream of CAP (Figs. 1,

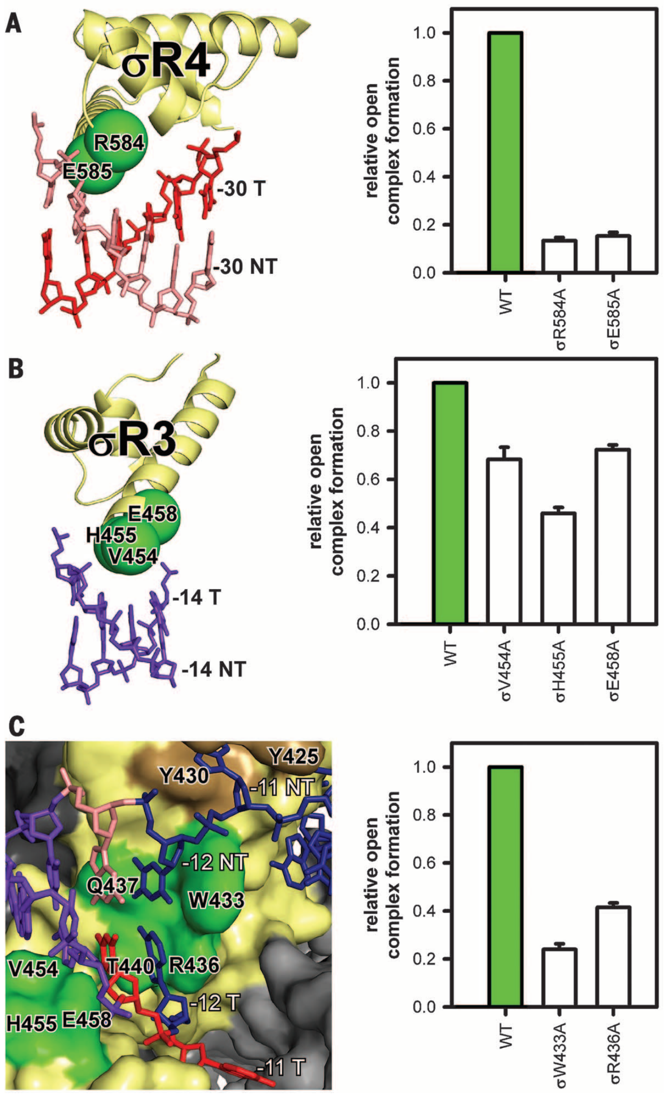


Fig. 2. Protein-DNA interactions that mediate promoter recognition. Green, σ residues that contact DNA bases (numbered as in *E. coli* σ^{70}); brown, σR2 residues that contact nontemplate-strand base –11. Other colors as in Fig. 1, B and C. Graphs, effects on RPo formation of Ala substitutions of *E. coli* σ^{70} (mean ± SEM; ≥3 determinations). (**A**) Interactions between σR4 and –35 region. (**B**) Interactions between σR3 and extended –10 element. (**C**) Interactions between σR2 and nontemplate- and template-strand nucleotides of first (–12NT, –12T) and second (–11NT, –11T) positions of –10 element.

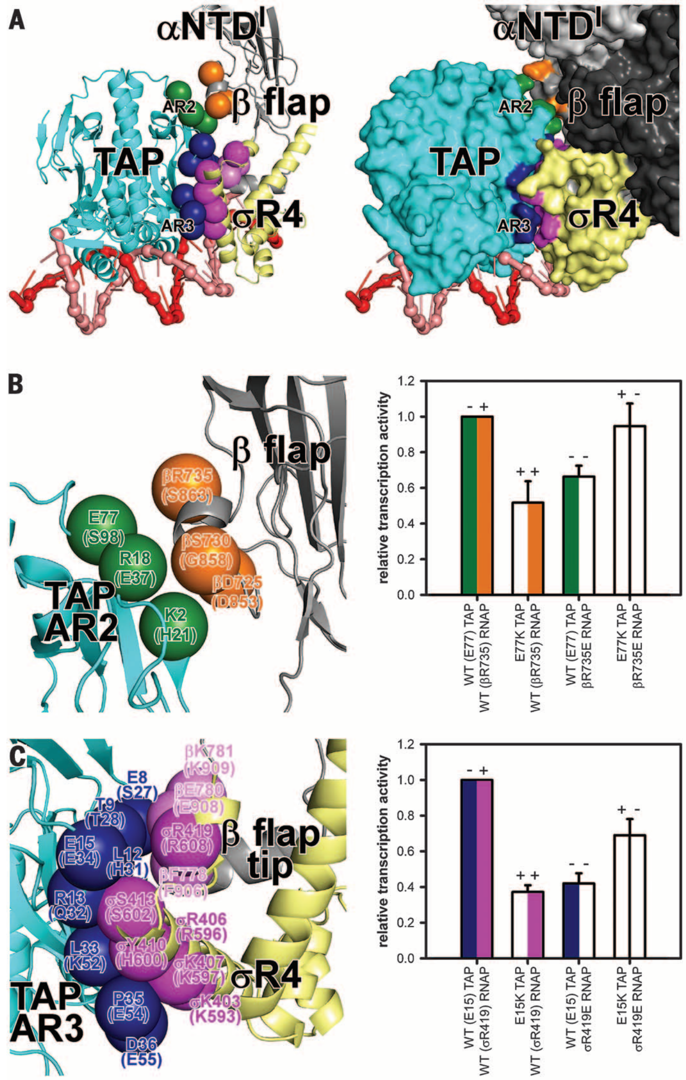
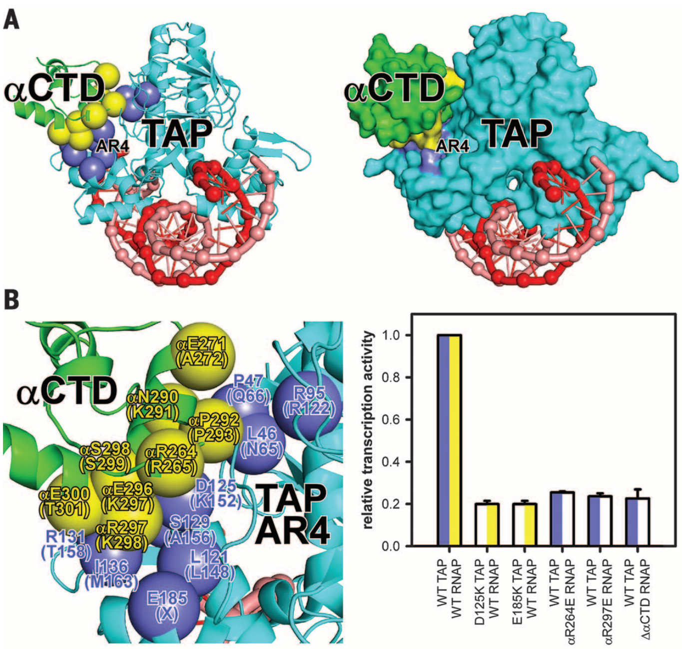


Fig. 3. Protein-protein interactions that mediate transcription activation: AR2 and AR3. Green, TAP AR2; blue, TAP AR3; orange, RNAP β-flap residues that contact AR2; magenta and light magenta, σR4 and RNAP β-flap-tip residues that contact AR3 (numbered as in TAP and *T. thermophilus* RNAP holoenzyme and, in parentheses, as in CAP and *E. coli* RNAP holoenzyme). Other colors as in Fig. 1, B and C. Graphs, effects on TAP-dependent transcription of single and double charge-reversal substitutions (mean ± SEM; ≥3 determinations). (**A**) Interactions between AR2, AR3, and RNAP holoenzyme (left, ribbons; right, surfaces). (**B**) AR2 interactions. (**C**) AR3 interactions.

B and C, and 4A) (*1*, *2*). Hydroxyl-radical DNA footprinting confirms that αCTD functions differently in *T. thermophilus* than in *E. coli.* Thus, *T. thermophilus* αCTD does not footprint DNA at a class II TAP-dependent promoter or a ribosomal RNA (rRNA) promoter (figs. S9 to S11), in contrast to *E. coli* αCTD, which footprints DNA immediately upstream of CAP at a class II CAP-dependent promoter and adenine-thymine–rich upstream-element DNA immediately upstream of the –35 element at a rRNA promoter (figs. S9 to S11). Consistent with the structure of TAP-RPo, fluorescence-polarization assays show that TAP is able to bind to αCTD in the absence of DNA and that the binding requires AR4 (Fig. 4C, left). Further consistent with the structure, fluorescence-polarization assays show that TAP is able to bind to RNAP holoenzyme in the absence of DNA and that the binding requires AR4 interactions and does not require AR2 and AR3 interactions (Fig. 4C, right).

The finding that TAP is able to bind to RNAP holoenzyme in the absence of DNA raises the possibility that TAP, in contrast to CAP, can access not only a "recruitment" pathway, in which the activator interacts first with DNA and then with RNAP holoenzyme, but also a "prerecruitment" pathway, in which the activator interacts first with RNAP holoenzyme and then with DNA (fig. S12) (*16*). Based on the equilibrium dissociation constant (K_D) for TAP-RNAP holoenzyme interaction (6 μM) (Fig. 4C, right) and the concentration of nontranscribing RNAP in bacteria in vivo (5 μM) (*17*), it appears likely that a prerecruitment pathway contributes to TAP-dependent transcription in *T. thermophilus* in vivo.

Measurements of effects of substitution of TAP-activating regions on the kinetics of transcription



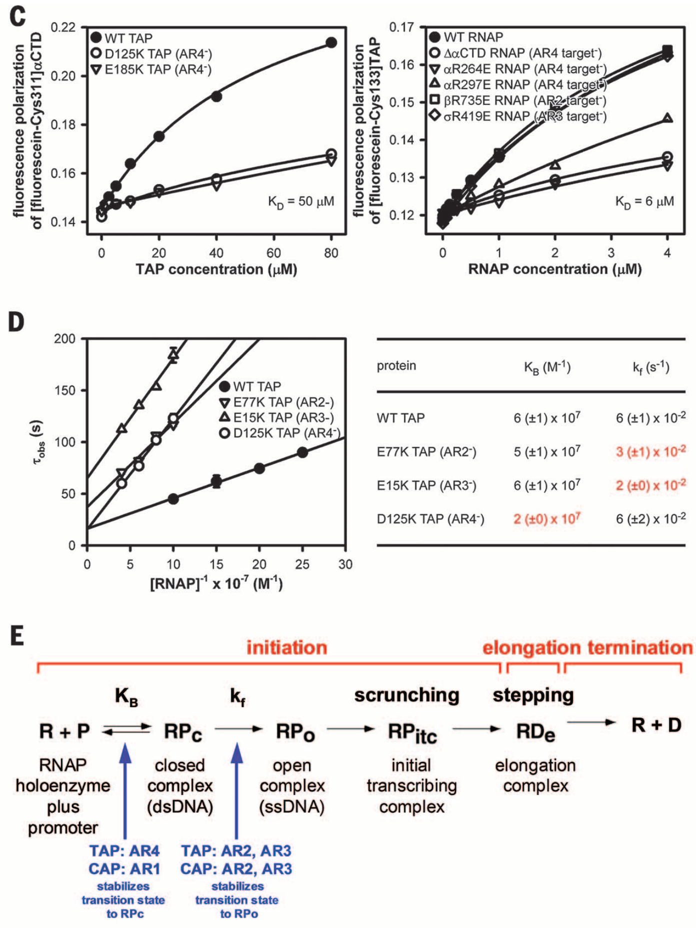


Fig. 4. Protein-protein interactions that mediate transcription activation: AR4 and kinetics. (**A** and **B**) TAP AR4 interaction with αCTD (left, ribbons; right, surfaces; lower left, close-up). Violet, TAP AR4; yellow, αCTD residues that contact AR4. Other colors as in Fig. 1, B and C. Graph, effects on TAP-dependent transcription of charge-reversal substitutions of AR4 and αCTD or truncation of αCTD (mean ± SEM; ≥3 determinations). (**C**) TAP-αCTD (left) and TAP-RNAP (right) interactions in the absence of DNA. (**D**) Data (left) and parameters (right) for effects of substitutions of AR2, AR3, and AR4 on kinetics of transcription initiation. (**E**) Summary of class II activator-dependent transcription.

initiation indicate that TAP AR2 and AR3 promote isomerization of RPc to RPo (k_f), and TAP AR4 promotes formation of RPc (K_B) (Fig. 4D). This pattern is reminiscent of the pattern for *E. coli* CAP, for which AR2 and AR3 promote isomerization of RPc to RPo (*11*, *13*, *15*), and AR1, through interaction with αCTD, promotes formation of RPc (*11*, *13*).

A long-standing question has been how a class II activator promotes isomerization of RPc to RPo, which entails loading of DNA into the RNAP active-center cleft, unwinding of DNA to form the transcription bubble, and closure of the RNAP clamp (*1*–*3*, *13*, *18*–*20*). The structure of TAP-RPo reveals that TAP does not interact with, and does not alter the conformation or interactions of, the RNAP active-center cleft, the transcription bubble, or the RNAP clamp. The structure further reveals that the interactions that promote isomerization—AR2 and AR3 interactions—are simple, adhesive, stabilizing protein-protein interactions between exposed surfaces of TAP and exposed surfaces of RNAP holoenzyme (Fig. 3 and fig. S8). We infer that interactions between a class II activator and RNAP holoenzyme that promote formation of RPc (AR4 interactions for TAP; AR1 interactions for CAP) and interactions between class II activator and RNAP holoenzyme that promote isomerization (AR2 and AR3 interactions) do not differ in character but, instead, differ only in timing (*13*, *18*–*20*). The former first occur in the transition state for formation of RPc and stabilize both RPc and RPo, whereas the latter first occur in the transition state for isomerization of RPc to RPo and stabilize RPo (Fig. 4E).

REFERENCES AND NOTES

1. S. Busby, R. H. Ebright, *J. Mol. Biol.* **293**, 199–213 (1999).
2. C. L. Lawson *et al.*, *Curr. Opin. Struct. Biol.* **14**, 10–20 (2004).
3. K. B. Decker, D. M. Hinton, *Annu. Rev. Microbiol.* **67**, 113–139 (2013).
4. B. P. Hudson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19830–19835 (2009).
5. Y. Agari, S. Kuramitsu, A. Shinkai, *Proteins* **80**, 1490–1494 (2012).
6. S. C. Schultz, G. C. Shields, T. A. Steitz, *Science* **253**, 1001–1007 (1991).
7. Y. Zhang *et al.*, *Science* **338**, 1076–1080 (2012).
8. Y. Zuo, T. A. Steitz, *Mol. Cell* **58**, 534–540 (2015).
9. B. Bae, A. Feklistov, A. Lass-Napiorkowska, R. Landick, S. A. Darst, *eLife* **4**, e08504 (2015).
10. B. Bae *et al.*, *eLife* **4**, e08505 (2015).
11. V. A. Rhodius, D. M. West, C. L. Webster, S. J. Busby, N. J. Savery, *Nucleic Acids Res.* **25**, 326–332 (1997).
12. C. L. Chan, M. A. Lonetto, C. A. Gross, *Structure* **4**, 1235–1238 (1996).
13. W. Niu, Y. Kim, G. Tau, T. Heyduk, R. H. Ebright, *Cell* **87**, 1123–1134 (1996).
14. M. A. Lonetto *et al.*, *J. Mol. Biol.* **284**, 1353–1365 (1998).
15. V. A. Rhodius, S. J. Busby, *J. Mol. Biol.* **299**, 311–324 (2000).
16. M. A. Zafar, I. M. Shah, R. E. Wolf Jr., *J. Mol. Biol.* **401**, 13–32 (2010).
17. M. Patrick, P. P. Dennis, M. Ehrenberg, H. Bremer, *Biochimie* **119**, 80–91 (2015).
18. M. Ptashne, A. Gann, *Nature* **386**, 569–577 (1997).
19. S. Roy, S. Garges, S. Adhya, *J. Biol. Chem.* **273**, 14059–14062 (1998).
20. S. L. Dove, F. W. Huang, A. Hochschild, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13215–13220 (2000).

ACKNOWLEDGMENTS

We thank the National Synchrotron Light Source for beamline access and E. Arnold for discussion. This work was funded by NIH grant GM041376 to R.H.E. The Protein Data Bank accession code is 5I2D.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1330/suppl/DC1
Materials and Methods
Figs. S1 to S12
Table S1
References (*21*–*55*)

8 February 2016; accepted 9 May 2016
10.1126/science.aaf4417

ENZYMOLOGY

Capture of a third Mg^{2+} is essential for catalyzing DNA synthesis

Yang Gao and Wei Yang*

It is generally assumed that an enzyme-substrate (ES) complex contains all components necessary for catalysis and that conversion to products occurs by rearrangement of atoms, protons, and electrons. However, we find that DNA synthesis does not occur in a fully assembled DNA polymerase–DNA–deoxynucleoside triphosphate complex with two canonical metal ions bound. Using time-resolved x-ray crystallography, we show that the phosphoryltransfer reaction takes place only after the ES complex captures a third divalent cation that is not coordinated by the enzyme. Binding of the third cation is incompatible with the basal ES complex and requires thermal activation of the ES for entry. It is likely that the third cation provides the ultimate boost over the energy barrier to catalysis of DNA synthesis.

Enzymes increase the rate of chemical reactions, which is thought to occur by a reduction in the activation energy required to reach the transition state (Fig. 1A) (*1–3*). Because of their transient and unstable nature, authentic transition states have not been visualized but are assumed to have the same chemical components as the substrate state. DNA polymerases, which catalyze a phosphoryltransfer reaction that incorporates deoxynucleoside triphosphates (dNTPs) into DNA, are known to require two Mg^{2+} ions (Fig. 1B) (*4–8*). Despite extensive kinetic studies using the stopped-flow technique and the dNTP analog dNTPαS, it remains controversial whether a conformational transition before catalysis (*9–14*) or the chemistry itself (*15*, *16*) is the rate-limiting step in DNA synthesis.

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA.
*Corresponding author. Email: wei.yang@nih.gov

We have recently visualized phosphodiester bond formation catalyzed by human DNA polymerase (Pol) η in crystallo (*17*). Consistent with the two–metal ion mechanism (*6–8*), binding of Mg^{2+} ions in the A and B sites occurs within 40 s and leads to alignment of the 3′-OH of the primer end with the α-phosphate of dNTP (Fig. 1C) (*17*). After another 40 s, product starts to appear without discernible conformational change of the enzyme or substrates. However, we observed a third Mg^{2+} ion appearing in a third "C" site after product formation (Fig. 1C) (*17*). An equivalent third metal ion, coordinated by the reaction products and four water molecules, has also been observed in the in-crystallo catalysis by DNA Pol β (*18–20*). Because of steric clashes with dNTP (Fig. 1C), the third metal ion cannot bind in Pol η enzyme-substrate (ES) complexes. Because of low occupancy in the C site and weak diffraction of Mg^{2+} ions, it has been unclear when the third Mg^{2+} appears and whether it is involved in the phosphoryltransfer reaction.

To determine the reaction coordinates of Pol η and the role of the third metal ion, we replaced Mg^{2+} with Mn^{2+}, which supports DNA synthesis (*21*) and is readily detected by x-ray diffraction even at low occupancy. Crystals of native Pol η (1 to 432 amino acids) in complex with DNA, deoxyadenosine triphosphate (dATP), and Ca^{2+} were grown at pH 6.0 in a nonreactive ground state (*17*). After exposure to a pH 7.0 reaction buffer containing 1 mM Mn^{2+} for 90 to 1800 s, crystals were flash frozen in liquid N_2 to stop the reaction, and 1.5 to 1.7 Å x-ray structures were determined at five reaction time points (table S1 and Materials and Methods). All five structures were practically identical, except for the gradual replacement of Ca^{2+} by Mn^{2+} in the B site (fig. S1). By 600 s, when ~90% of the A and B sites were occupied by Mn^{2+}, the 3′-OH of the DNA primer was aligned with the α-phosphorus of the dATP, and the structure was identical to that of crystals soaked in 1 mM Mg^{2+} for 40 s (fig. S2, A and B). Similar to the reaction in Mg^{2+} (*17*), a water molecule (WatN) hydrogen bonded to the nucleophilic 3′-OH appeared at 90 s, and its occupancy increased with time in correlation with binding of the A-site Mn^{2+} (fig. S2, C and D). However, in 1 mM Mn^{2+}, the Pol η–DNA complex remained in the substrate state with no product and no C-site Mn^{2+} ion for up to 1800 s (Fig. 1D).

We assayed the metal ion (Me^{2+}) requirements for Pol η catalysis in solution and found that 0.6 mM Mg^{2+} or 2.7 mM Mn^{2+} is needed to attain the half-maximal reaction rate (Fig. 2A and table S2). We then examined the Mn^{2+} affinity of each binding site in crystallo. Although increasing the Mn^{2+} concentration (0.5 to 15 mM) accelerated the rate of metal-ion binding in all three sites (Fig. 2B and table S3), the apparent dissociation constant (K_d) values of the A and

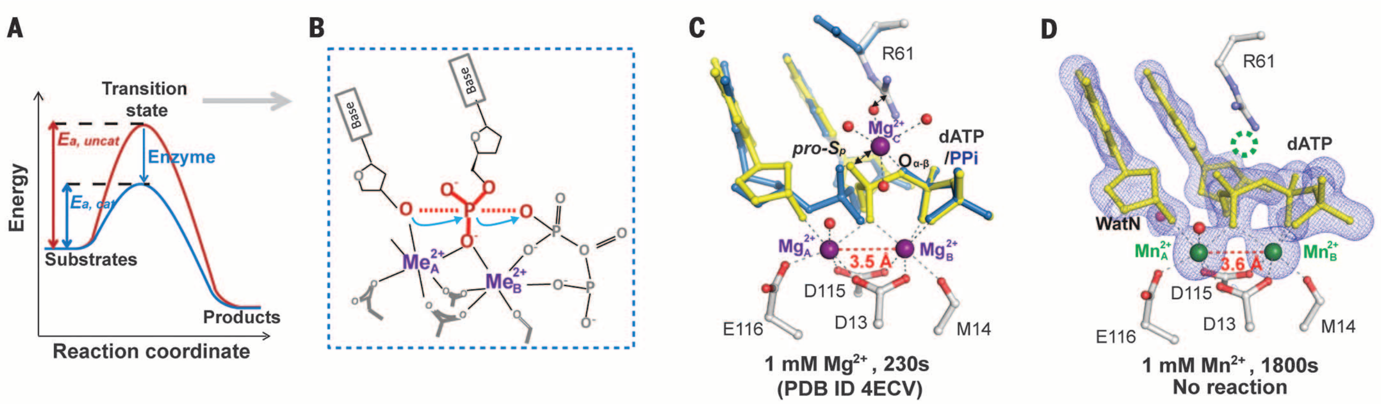


Fig. 1. No DNA synthesis without the third metal ion. (**A**) Reaction coordinate of enzyme catalysis. (**B**) The assumed transition state of the two–metal ion catalysis. (**C**) The structure of Pol η catalyzing DNA synthesis in crystallo (PDB: 4ECV) (*17*). The C-site Mg^{2+} is coordinated by the products (60%, blue) but clashes with the substrate dATP and R61 side chain (40%, yellow). (**D**) The structure of Pol η incubated with substrates and 1 mM Mn^{2+} for 1800 s. No C-site metal ion or reaction products were detected. The corresponding $F_{obs} - F_{calc}$ (F_o-F_c) omit map contoured at 3.5 σ (blue mesh) is superimposed.

B sites were below 0.5 mM. The K_d for the C site, however, was 3.2 mM, close to the 2.7 mM measured in solution (Fig. 2A).

When in-crystallo reactions occurred in 10 mM Mn^{2+}, catalysis proceeded as in 1 mM Mg^{2+} (*17*), except that the A-site Mn^{2+} did not dissociate upon product formation as does Mg^{2+} (fig. S3, A and B) and slightly less product accumulated with Mn^{2+} than with Mg^{2+}. However, unlike the reaction in Mg^{2+}, the C-site Mn^{2+} appeared simultaneously with the reaction product, 30 s after binding of the two canonical metal ions (Fig. 2C). Electron density for the new phosphodiester bond and the C-site Mn^{2+}, whose chemical nature was confirmed by its anomalous diffraction and characteristic octahedral coordination geometry (fig. S3C), had one-to-one correlation at every time point and Mn^{2+} concentration (Fig. 2D). In Mg^{2+} by contrast, with 15% product formed at 80 s, the C-site Mg^{2+} was at too low occupancy to be observed and was not detected until 140 s when product had accumulated to 40% (fig. S3A) (*17*). Previous stopped-flow studies indicate that one of the metal ion–binding sites has much lower affinity for Mg^{2+} and, thus, limits DNA synthesis (*16*). Our in-crystallo titrations unequivocally show that the low-affinity binding site is neither A nor B but the C site, which determines the concentration of Mg^{2+} or Mn^{2+} necessary for the DNA synthesis reaction.

The C-site Me^{2+} is coordinated by four water molecules and two oxygen atoms, one each from the product DNA and pyrophosphate, which correspond to the α pro-S_p oxygen and the α,β bridging oxygen of dNTP (Fig. 1C). Sulfur substitution of the pro-S_p oxygen (S_p-dNTPαS) has been widely used to dissect the reaction kinetics of DNA synthesis (*11–14*, *22*), because the pro-S_p atom is not directly involved in A- or B-site Me^{2+} coordination. The reduction of the reaction rate by S_p-dNTPαS has been interpreted to be "conformational" (for reduction of less than three-quarters) or to affect the chemistry itself (for a reduction of more than three quarters) (*11–14*, *23*).

As a ligand of the third Me^{2+}, the sulfur in S_p-dATPαS was tolerated by Pol η (table S2) but required much higher [Mg^{2+}] (15 mM) and [Mn^{2+}] (9 mM) than dATP for catalysis to occur (Figs. 2A and 3A). Unexpectedly, in-crystallo S_p-dATPαS slowed Mg^{2+} and Mn^{2+} binding at the A site. After a lengthened delay, product started to form, but the C site remained empty (Fig. 3B and fig. S4, A to C). We suspect that the third Mn^{2+} still assisted product formation, but the association was too transient to be observed. In addition, although A-site Mg^{2+} occupancy was reduced in the presence of dATPαS, an alternative A′ site appeared 2.6 Å away (fig. S4, D and E). These data suggest that the reduced reaction rate with S_p-dATPαS cannot be attributed to conformational effects (*11–14*) but involves impaired A- and C-site Mg^{2+} binding and altered reaction chemistry.

To bind the third Me^{2+}, the arginine 61 (R61) side chain, which forms salt bridges with the dNTP (*17*), moves to vacate the C site (Figs. 1C and 3B). When alanine replaces arginine at position 61 (R61A), the enzymatic rate (k_{cat}) is reduced by two-thirds (table S2) (*24*, *25*), but the metal-ion requirement and the general reaction process in crystallo were indistinguishable from wild-type (WT) Pol η (fig. S5, E and F). However, the delay between binding of two Mg^{2+} ions and product formation was lengthened from WT's 40 s to R61A's 160 s (Fig. 3C). This delay likely stems from a slight shift of dATP away from the active site and a 0.3 Å increase in separation between the 3′-OH and α-phosphate (Fig. 3D). The void left by the R61A mutation was occupied by water molecules (*25*) and not by the abundant K^+ or Rb^+ (identifiable by anomalous diffraction) in the reaction buffer (fig. S5). The subtle misalignment of the substrate, which was repeatedly observed with R61A and R61M (in which methionine replaces Arg[61]) mutant Pol η (*25*) and with dATPαS, led to a prolonged delay before C-site Me^{2+} binding and product formation (fig. S4, B and C).

Notably, a +1 charged side chain at the position equivalent to R61 is found in all A-, B-, and Y-family DNA polymerases and reverse transcriptases, despite diverse structures of finger domains surrounding the C site (fig. S6). Among C- and X-family DNA polymerases, there is no R61 equivalent, but the third metal ion has been observed for the X-family Pol β (*18–20*). The finger domains, which carry the +1 charged residue, distinguish correct from incorrect incoming nucleotides by enclosing only a correct dNTP (*26*). A closed finger appears to be a prerequisite for C-site metal-ion binding and catalysis. The varied environment surrounding the C site may thus be exploited for drug design to increase specificity and to reduce toxicity of broadly used nucleoside and nucleotide analogs targeting DNA polymerases in antiviral and anticancer therapeutics (*27*, *28*).

Because the C site does not exist in the Pol η–substrate complex but is required for product formation, we hypothesized that thermal motion of the well-aligned reactants in the ES complex may create an opening for the third metal ion. If so, elevated temperature would promote C-site metal-ion binding and thus catalysis. To test this hypothesis, we designed a two-step in-crystallo reaction. The Pol η crystals were first soaked in 1 mM Mn^{2+} to saturate the A and B sites and then exposed to 5 mM Mn^{2+} at 4°C to 37°C for 60 s for catalysis to occur (Fig. 4A). The diffusion rate of Mn^{2+} in crystallo was unaffected by temperature, as demonstrated by Mn^{2+} binding at the A site (Fig. 4B). But in the two-step reaction, no C-site Mn^{2+} or product was detected at 4°C (Fig. 4C). At 14°C, low levels of the third Mn^{2+} ion and products were observed, and their

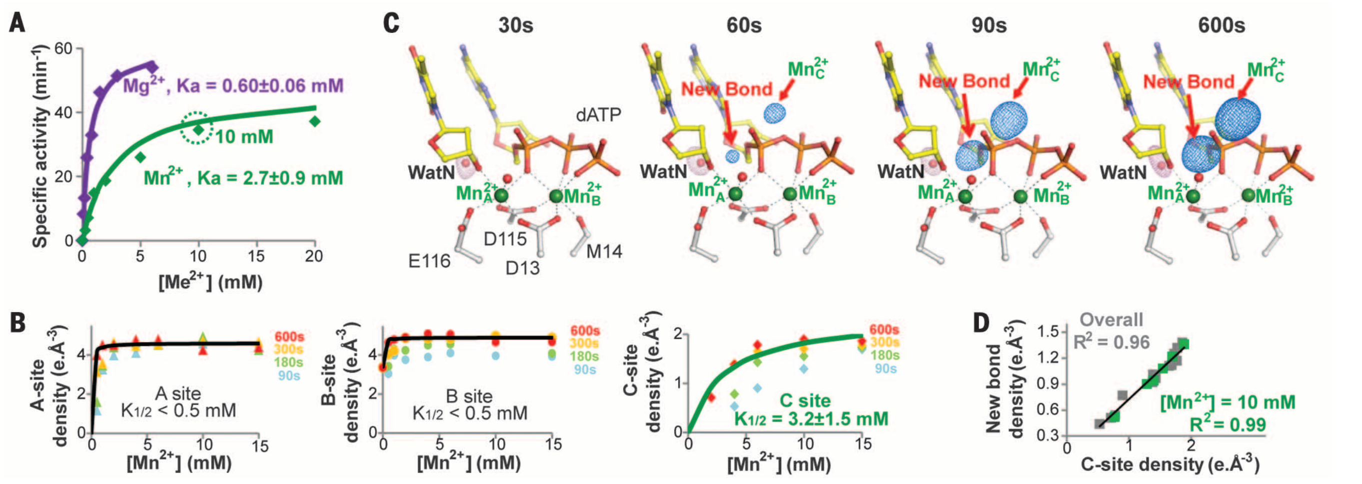


Fig. 2. Coupled appearance of the third Mn^{2+} and reaction products. (**A**) Mg^{2+} (purple) and Mn^{2+} (green) dependence of Pol η catalysis in solution. K_a, activation constant. (**B**) Titration of the A-, B-, and C-site Mn^{2+} binding in crystallo. The 600-s data were fitted to equilibrium binding modes to yield the K_d values. (**C**) Structures of Pol η in-crystallo catalysis with 10 mM Mn^{2+}. The F_o-F_c omit map for the new bond, the C-site Mn^{2+} (blue) and the WatN (pink) were contoured at 3σ and superimposed onto each structure. (**D**) Correlation (R^2) between the new bond formation and the C-site Mn^{2+} binding.

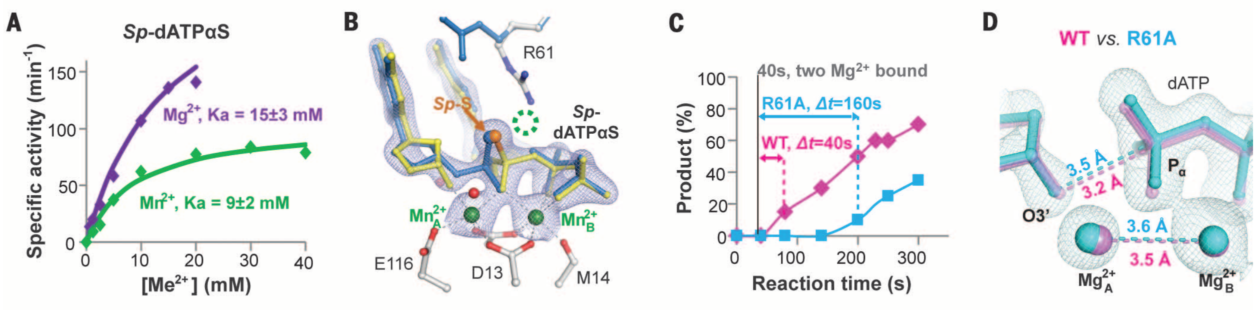


Fig. 3. Changing the C-site environment alters Pol η catalysis. (**A**) Mg^{2+} (purple) and Mn^{2+} (green) dependence of Pol η incorporating dATPαS in solution. (**B**) In crystallo incorporation of dATPαS by Pol η with 20 mM Mn^{2+} at 600 s showed product formation (50%) but no C-site Mn^{2+}. The $2F_o$-F_c map contoured at 2σ level (blue meshes) is superimposed. (**C**) Time delay in product formation by WT (magenta) and R61A (cyan) Pol η in crystallo. (**D**) Deviation of dATP in the ES of R61A Pol η [cyan with $2F_o$-F_c map contoured at 1.5σ] from WT Pol η (magenta).

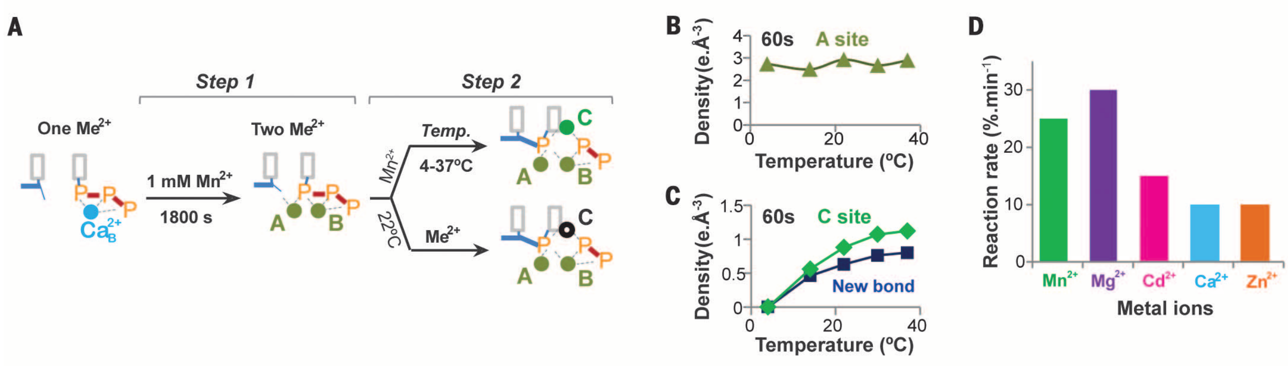


Fig. 4. Thermal energy-dependent C-site formation and its metal-ion selectivity. (**A**) Schematic diagram of the two-step in-crystallo reactions that probe the C-site formation and ion selectivity. (**B**) Binding of the A-site Mn^{2+} was unaffected by varying temperature from 4°C to 37°C. (**C**) Binding of the C-site Mn^{2+} and the product formation increased with the temperature from 4°C to 37°C. (**D**) Rates of product formation with five metal ions tested in the second step.

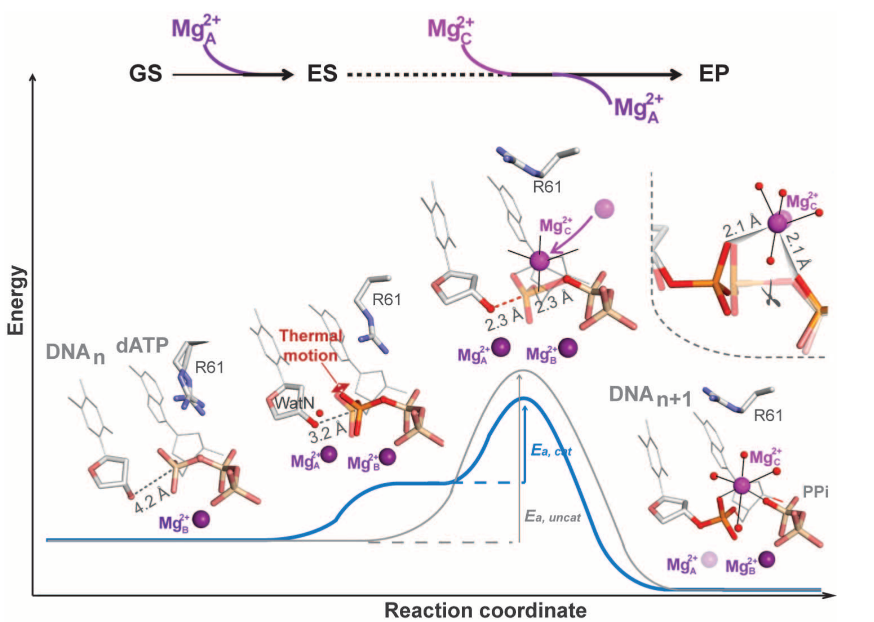


Fig. 5. Mechanism of Pol η catalysis. Pol η binds DNA and an incoming dNTP along with the B-site Mg^{2+} to form the ground state (GS). Binding of the A-site Mg^{2+} leads to the reaction-ready ES state, in which the 3′-OH is aligned with dNTP and WatN is recruited. Thermal motions of the reactants create the C site, which leads to the third metal-ion binding and the transition state (TS) formation. The C-site metal ion promotes the phosphoryltransfer from the leaving group to the nucleophilic 3′-OH by which it overcomes the energy barrier to the product state (PS). E_a, the activation energy of the reaction, catalyzed or uncatalyzed.

amounts doubled at 30°C. The temperature dependence of C-site and product formation corroborates that binding of the third metal ion is rate limiting in the DNA synthesis reaction.

To determine the metal-ion selectivity at the C site, we varied Me^{2+} in the second step of the two-step reaction (Fig. 4A). Catalysis occurred most efficiently with Mg^{2+}, followed by Mn^{2+} and Cd^{2+} (Fig. 4D and fig. S7). Ca^{2+} and Zn^{2+} seemingly also led to product formation at ~40% efficiency of Mg^{2+}. However, the C-site coordination geometry with all five Me^{2+} tested appeared identical to that of Mg^{2+} and Mn^{2+} (fig. S7), despite different coordination distances of Ca^{2+} (2.3 to 2.5 Å) and Mg^{2+} or Mn^{2+} (2.1 Å). It is thus likely that Ca^{2+} and Zn^{2+} replaced A- or B-site Mn^{2+} in some Pol η molecules and that the freed Mn^{2+} ions may occupy the C site in other Pol η molecules to support the catalysis. The low affinity and strong preference for Mg^{2+} at the C site, which cannot be explained by its coordination ligands, suggest a catalytic role for the third metal ion in DNA synthesis.

On the basis of the requirement for three metal ions in DNA synthesis, we suggest a revision of the catalytic mechanism (Fig. 5). DNA synthesis begins with binding of dNTP along with the B-site Mg^{2+} and formation of a ground-state Pol η–DNA-dNTP-Mg^{2+} complex (GS). Watson-Crick pairing between the template and dNTP

favors A-site Mg^{2+} binding. The two Mg^{2+} ions and the R61 side chain neutralize and align dNTP with DNA in the reaction-ready state (ES), where the juxtaposed and polarized substrates recruit WatN (fig. S2) (*17*). However, neither deprotonation nor chemistry takes place without the C-site Mg^{2+}. Thermal motion may transiently bring the perfectly aligned reactants closer to each other by fractions of an angstrom and may create an entry for the third Mg^{2+}. Close approach of the reactants may also increase negative charge around the α-phosphate and favor replacement of the +1 charged R61 by the C-site Mg^{2+}. We hypothesize that the energy barrier to the transition state is overcome by binding of the third Mg^{2+}. The stringent octahedral coordination geometry of Mg^{2+} implies that the C-site Mg^{2+} may help to break the α-β phosphodiester bond (Fig. 5) in addition to protonating the pyrophosphate leaving group (*17*). Product formation is coupled to disappearance of WatN (fig. S3D), which likely deprotonates the 3′-OH, and to release of the A-site Mg^{2+}, which prevents the reverse reaction (fig. S3, A and B).

It has long been assumed that enzymes stabilize transition states and reduce the energy barrier to product formation (Fig. 1A), but de novo design of enzymes based on this assumption has not been successful (*29*–*32*). Notwithstanding its crucial role in catalysis, the C-site metal ion of polymerases has evaded detection by biochemical and structural studies of DNA polymerases for decades. Identification of the essential third metal ion in the Pol η catalysis leads us to anticipate that acquisition of transient metal-ion cofactors in transition states may be a general feature that enables enzyme catalysis.

REFERENCES AND NOTES

1. M. Garcia-Viloca, J. Gao, M. Karplus, D. G. Truhlar, *Science* **303**, 186–195 (2004).
2. S. J. Benkovic, S. Hammes-Schiffer, *Science* **301**, 1196–1202 (2003).
3. A. Warshel *et al.*, *Chem. Rev.* **106**, 3210–3235 (2006).
4. P. J. Rothwell, G. Waksman, *Adv. Protein Chem.* **71**, 401–440 (2005).
5. M. D. Tsai, *Biochemistry* **53**, 2749–2751 (2014).
6. T. A. Steitz, J. A. Steitz, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6498–6502 (1993).
7. T. A. Steitz, *J. Biol. Chem.* **274**, 17395–17398 (1999).
8. W. Yang, J. Y. Lee, M. Nowotny, *Mol. Cell* **22**, 5–13 (2006).
9. C. M. Joyce, S. J. Benkovic, *Biochemistry* **43**, 14317–14324 (2004).
10. K. A. Johnson, *J. Biol. Chem.* **283**, 26297–26301 (2008).
11. S. S. Patel, I. Wong, K. A. Johnson, *Biochemistry* **30**, 511–525 (1991).
12. M. E. Dahlberg, S. J. Benkovic, *Biochemistry* **30**, 4835–4843 (1991).
13. M. T. Washington, L. Prakash, S. Prakash, *Cell* **107**, 917–927 (2001).
14. K. A. Fiala, Z. Suo, *Biochemistry* **43**, 2116–2125 (2004).
15. P. J. Rothwell, V. Mitaksov, G. Waksman, *Mol. Cell* **19**, 345–355 (2005).
16. M. Bakhtina *et al.*, *Biochemistry* **44**, 5177–5187 (2005).
17. T. Nakamura, Y. Zhao, Y. Yamagata, Y. J. Hua, W. Yang, *Nature* **487**, 196–201 (2012).
18. B. D. Freudenthal, W. A. Beard, D. D. Shock, S. H. Wilson, *Cell* **154**, 157–168 (2013).
19. B. D. Freudenthal *et al.*, *Nature* **517**, 635–639 (2015).
20. R. Vyas, A. J. Reed, E. J. Tokarsky, Z. Suo, *J. Am. Chem. Soc.* **137**, 5225–5230 (2015).
21. J. A. Cowan, *Chem. Rev.* **98**, 1067–1088 (1998).
22. P. M. Burgers, F. Eckstein, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4798–4800 (1978).
23. D. Herschlag, J. A. Piccirilli, T. R. Cech, *Biochemistry* **30**, 4844–4854 (1991).
24. C. Biertümpfel *et al.*, *Nature* **465**, 1044–1048 (2010).
25. Y. Su, A. Patra, J. M. Harp, M. Egli, F. P. Guengerich, *J. Biol. Chem.* **290**, 15921–15933 (2015).
26. E. Y. Wu, L. S. Beese, *J. Biol. Chem.* **286**, 19758–19767 (2011).
27. T. N. Kakuda, *Clin. Ther.* **22**, 685–708 (2000).
28. L. P. Jordheim, D. Durantel, F. Zoulim, C. Dumontet, *Nat. Rev. Drug Discov.* **12**, 447–464 (2013).
29. J. Wagner, R. A. Lerner, C. F. Barbas 3rd, *Science* **270**, 1797–1800 (1995).
30. D. N. Bolon, S. L. Mayo, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14274–14279 (2001).
31. L. Jiang *et al.*, *Science* **319**, 1387–1391 (2008).
32. D. Röthlisberger *et al.*, *Nature* **453**, 190–195 (2008).

ACKNOWLEDGMENTS

Y.G. carried out all experiments; W.Y. conceived and designed the project, and both authors interpreted data and prepared the manuscript. We thank D. J. Leahy and M. Gellert for critical reading and editing the manuscript. This work is funded by NIH intramural program (DK036146-08, W.Y.).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1334/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S3
References (*33*–*44*)

8 December 2015; accepted 10 May 2016
10.1126/science.aad9633

CANCER IMMUNOTHERAPY

Targeting of cancer neoantigens with donor-derived T cell receptor repertoires

Erlend Strønen,[1,2] Mireille Toebes,[3] Sander Kelderman,[3] Marit M. van Buuren,[3] Weiwen Yang,[1,2] Nienke van Rooij,[3] Marco Donia,[4] Maxi-Lu Böschen,[1,2] Fridtjof Lund-Johansen,[2,5] Johanna Olweus,[1,2]*† Ton N. Schumacher[3]*†

Accumulating evidence suggests that clinically efficacious cancer immunotherapies are driven by T cell reactivity against DNA mutation–derived neoantigens. However, among the large number of predicted neoantigens, only a minority is recognized by autologous patient T cells, and strategies to broaden neoantigen-specific T cell responses are therefore attractive. We found that naïve T cell repertoires of healthy blood donors provide a source of neoantigen-specific T cells, responding to 11 of 57 predicted human leukocyte antigen (HLA)–A*02:01–binding epitopes from three patients. Many of the T cell reactivities involved epitopes that in vivo were neglected by patient autologous tumor-infiltrating lymphocytes. Finally, T cells redirected with T cell receptors identified from donor-derived T cells efficiently recognized patient-derived melanoma cells harboring the relevant mutations, providing a rationale for the use of such "outsourced" immune responses in cancer immunotherapy.

Accumulating data suggest that tumor regression induced by cancer immunotherapies that exploit the endogenous T cell pool (*1*, *2*) relies on recognition of neoantigens that are formed as a consequence of tumor-specific DNA mutations. A striking observation in cancer patients and in mouse models is that neoantigen-specific T cell reactivity is generally limited to just a few mutant epitopes, even though the number of predicted epitopes is large (*3*–*12*). This scarcity of T cell–recognized neoantigens could potentially reflect immune editing of tumors by T cells (*13*). Alternatively, an effector T cell pool toward many tumor-expressed neoantigens may be absent because of ineffective priming or because of tolerization of these T cells. Recent work has shown that vaccination with neoantigen peptide–loaded dendritic cells can increase the breadth of mutant peptide-specific T cells in melanoma patients (*14*). In that study, it could not be established whether newly induced T cells could recognize autologous tumor cells. Nonetheless, these data provide a further incentive for the development of strategies that broaden neoantigen-specific T cell reactivity.

Here, we aimed to establish whether T cell receptors (TCRs) that are obtained outside of the autologous T cell repertoire can be used to engineer neoantigen-specific T cell immunity. To this end, we generated immune responses to HLA-A*02:01-restricted neoantigens from the non-tolerized T cell repertoires derived from donors that express this allele. Using this approach, we evaluated (i) whether donor-derived T cells can recognize relevant tumor cells, (ii) whether such "outsourced" immune responses provide evidence

[1]Department of Cancer Immunology, Oslo University Hospital Radiumhospitalet, Oslo, Norway. [2]K. G. Jebsen Centers for Cancer Immunotherapy and for Inflammation Research, Institute for Clinical Medicine, University of Oslo, Oslo, Norway. [3]Division of Immunology, Netherlands Cancer Institute, Amsterdam, Netherlands. [4]Center for Cancer Immune Therapy, Department of Hematology, Herlev Hospital, University of Copenhagen, Herlev, Denmark. [5]Department of Immunology and Transfusion Medicine, Oslo University Hospital Rikshospitalet, Oslo, Norway.
***Corresponding author. Email: johanna.olweus@medisin.uio.no (J.O.); t.schumacher@nki.nl (T.N.S.)** †These authors contributed equally to this work.

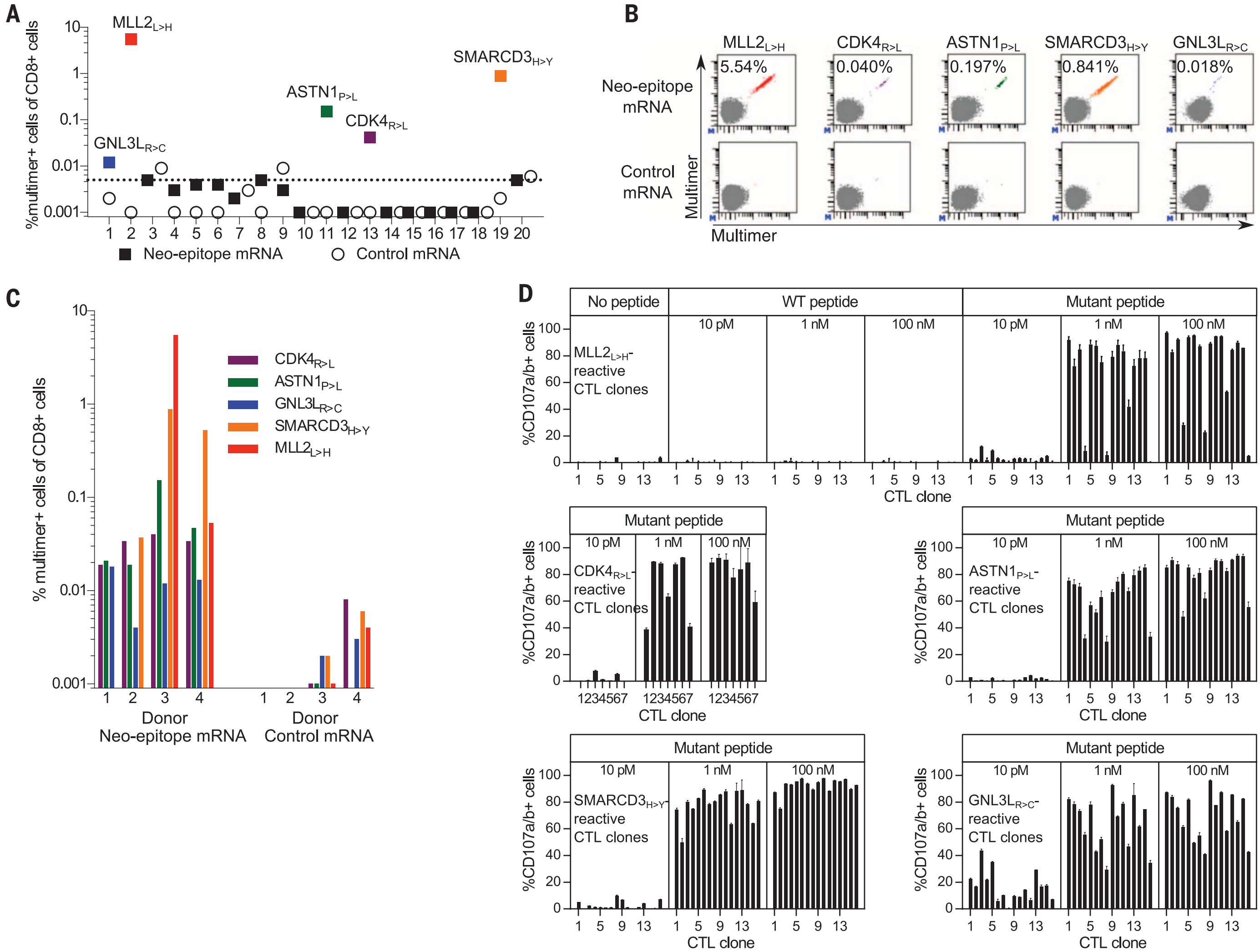
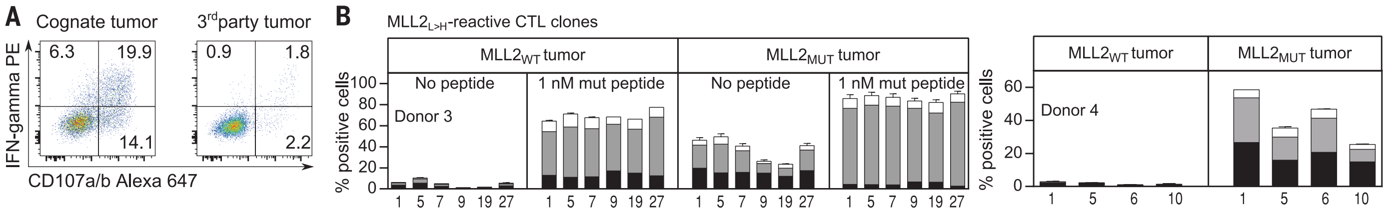


Fig. 1. In vitro induction and functional activity of donor-derived neoantigen-reactive T cells. Data depict T cell responses against predicted HLA-A*02:01-binding neoantigens from patient 1. (**A**) PBMCs (donor 3) stimulated with autologous antigen-presenting cells (APCs) transfected with mRNA encoding either predicted neoantigens (solid squares) or CT/CD20 control antigens (open circles) were stained with pMHC multimers complexed with predicted epitopes. Symbols indicate percentage of live CD8^{pos} cells staining positively for pMHC multimers complexed with indicated peptides. Colored squares indicate populations sorted for further analysis. (**B**) Flow cytometry analysis. (**C**) Magnitude of multimer^{pos} T cell populations for the indicated predicted neoantigens induced by APCs transfected with mRNA encoding either relevant neoepitopes (left) or control CT/CD20 epitopes (right) from four healthy donors. (**D**) Degranulation responses of CTL clones (donor 4) analyzed as shown in fig. S2. Each graph represents the reactivity of 7 to 16 clones to the indicated neoantigen. Controls are depicted only for MLL2$_{L>H}$-reactive CTL clones; corresponding data for remaining clones are shown in fig. S3A. Graphs are representative of CTL clones from all donors tested and show means of triplicates. Error bars denote SD.

for a neglected pool of neoantigens on human cancers, and (iii) which types of mutant peptides are best detected by the T cell–based immune system.

To determine the feasibility of using donor-derived T cell pools to induce neoantigen-specific T cell reactivity, we initially focused on an HLA-A*02:01^{pos} stage IV melanoma patient. Whole-exome and RNA sequencing of tumor material revealed 249 nonsynonymous mutations within expressed genes, and 126 mutant epitopes were predicted to bind to HLA-A*02:01 (*15*). Of these 126 neopeptides, only two were detected by T cells grown from the same tumor lesion. To investigate whether a larger fraction of predicted neoepitopes could be recognized by a healthy donor immune system, we selected 20 candidate neoepitopes based on high predicted binding affinity to HLA-A*02:01 (table S1). Nonadherent peripheral blood mononuclear cells (PBMCs) from healthy donors were then cocultured with autologous monocyte-derived dendritic cells transfected with mRNA encoding the candidate epitopes in a tandem minigene configuration, or with a control minigene encoding known epitopes from cancer/testis (C/T) antigens and CD20 (*16*) that were recognized by relevant cytotoxic T lymphocytes (CTLs) (fig. S1). Analyses of resulting cell populations by peptide–major histocompatibility complex (pMHC) multimer staining revealed T cell reactivity toward 5 of 20 neoantigens from patient 1, whereas such reactivity was negligible in control cultures (Fig. 1, A and B). Analysis of T cell reactivity from three additional donors revealed three to five neoantigen-specific T cell responses in all cases (Fig. 1C). One of the T cell responses reproducibly induced in this system, to the neoantigen CDK4$_{R>L}$, was also one of two responses detected among tumor-infiltrating lymphocytes (TILs) of patient 1.

pMHC-multimer^{pos} CD8 cells were sorted from donors 2, 3, and 4 to generate CTL clones. Resulting clones that stained positively with relevant pMHC multimers (>82% of clones) were then tested for functional activity using a live-cell barcoding assay (fig. S2). Analysis of 185 CTL clones revealed reactivity of the majority of clones toward target cells pulsed with mutant peptide



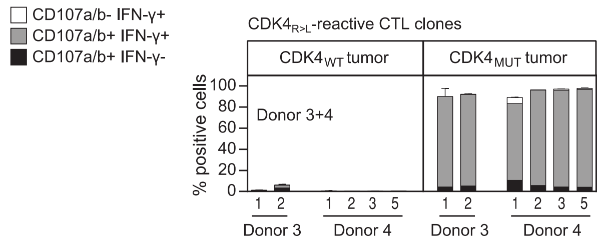
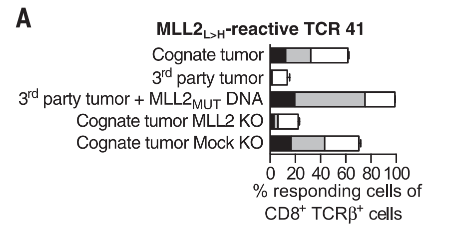


Fig. 2. Tumor recognition by donor-derived neoantigen-reactive CTLs. (**A**) $MLL2_{L>H}$-reactive CTL clone 7 (donor 3) was incubated with a melanoma line derived from patient 1 that carried the mutated MLL2 gene (cognate tumor; left) or with a third-party HLA-A*02:01^{pos} melanoma line carrying the wild-type MLL2 gene (right). Data depict the percentage of IFN-γ^{pos} and/or CD107a/b^{pos} cells of live CD8 cells. (**B**) Neoantigen-reactive CTL clones (donors 3 and 4) incubated with the indicated melanoma lines were analyzed as described in (A). Where indicated, melanoma lines were pulsed with 1 nM corresponding neoantigen (shown for $MLL2_{L>H}$-reactive CTL clones from donor 3). All $MLL2_{L>H}$- and $CDK4_{R>L}$-reactive CTL clones were selected for TCR sequencing. Graphs show means of duplicates; error bars denote SD.



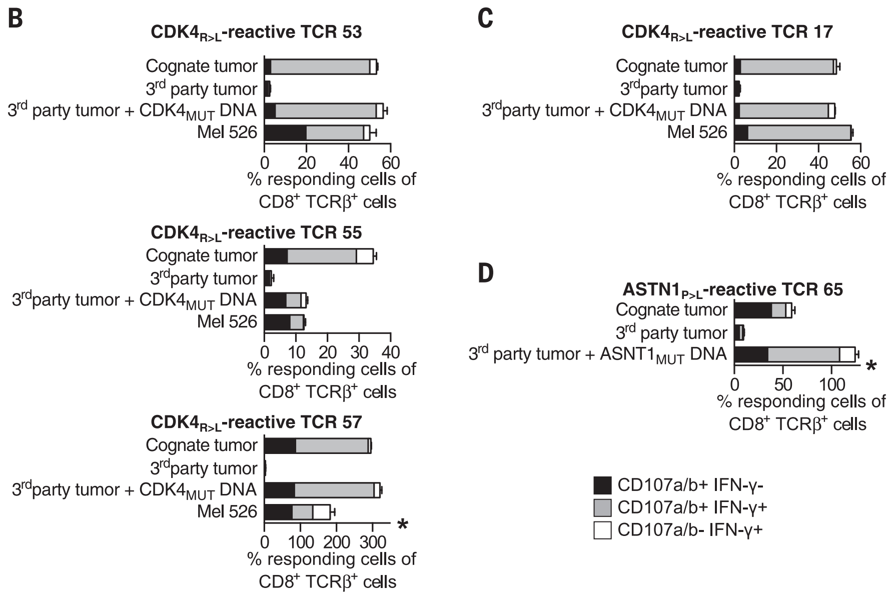


Fig. 3. Tumor recognition by genetic transfer of donor-derived neoantigen-specific TCRs. Healthy donor peripheral blood T cells transduced with indicated donor-derived TCRs [(A), (B), (D)] or a patient-derived TCR (C) were incubated with the cognate melanoma line, an HLA-A*02:01^{pos} third-party melanoma line, or melanoma cells modified as indicated. Percentage of responding CD8 cells was analyzed as described in Fig. 2. (**A**) Melanoma cells in which the mutant MLL2 gene was knocked out (cognate tumor MLL2 KO), mock-treated cognate melanoma cells (cognate tumor Mock KO), or third-party melanoma cells stably transfected with DNA encoding the relevant mutant neoantigen (3rd party tumor + $MLL2_{MUT}$ DNA) were used as target cells. (**B** to **D**) Indicated melanoma cells were used as target cells for indicated healthy donor–derived [(B) and (D)] and patient-derived (C) TCRs. Data for each TCR are representative of two or three independent experiments using T cells from different healthy donors. Graphs depict mean of duplicate samples; error bars denote SD. Values were corrected for transduction efficiency, measured as percentage of CD8^{pos} cells staining positively with antibody to mouse TCRβ chain. Asterisk indicates TCRs for which the fraction of TCR-expressing T cells is underestimated by staining with antibody to mouse TCRβ chain constant domain (fig. S6).

at concentrations down to 1 nM and below, with negligible recognition of the wild-type counterpart (Fig. 1D and figs. S2B and S3).

To assess recognition of a short-term melanoma line of patient 1, we selected 76 CTL clones that specifically recognized target cells pulsed with the corresponding neoantigens at low concentrations. All $MLL2_{L>H}$-reactive CTL clones tested (n = 10) recognized the relevant melanoma cells. In contrast, no recognition of an HLA-A*02:01^{pos} third-party melanoma was observed, unless pulsed with $MLL2_{L>H}$ peptide (Fig. 2, A and B). By the same token, all $CDK4_{R>L}$-reactive CTL clones tested (n = 6) showed vigorous and specific reactivity toward CDK4 mutant melanoma cells (Fig. 2B). Among CTL clones reactive with $ASTN1_{P>L}$ and $SMARCD3_{H>Y}$, 7 of 24 clones and 5 of 20 clones, respectively, showed recognition of cognate melanoma (fig. S4). None of the $GNL3L_{R>C}$-reactive CTL clones tested (n = 16) recognized cognate or third-party melanoma unless pulsed with the relevant neoantigen (fig. S4).

T cell inductions were subsequently performed for predicted neoantigens from tumors of two additional patients. For patient 2, a set of 27 neopeptides (table S2) with a median predicted binding affinity to HLA-A*02:01 of 34 nM (range 2 to 140 nM) was selected from among 154 mutant peptides predicted to bind to HLA-A*02:01. No HLA-A*02:01–restricted neoantigen-specific T cell responses had been detected in TILs isolated from this patient when screening for reactivity to these 154 peptides. In contrast, responses to six predicted neoantigens were induced among T cells derived from four healthy donors (fig. S5, A and B). For patient 3, no T cell responses to 10 predicted neoantigens were detected (table S3). Predicted binding affinities of these potential neoantigens were considerably lower (median 225 nM) than those from patient 1 (median 41 nM) and patient 2 (median 34 nM).

From pMHC-multimer^{pos} CD8 cells from donors 5, 7, and 8, we established CTL lines reactive with the $USP28_{C>F}$, $SNX24_{P>L}$, $PGM5_{H>Y\text{-}462\text{-}470}$, and $PGM5_{H>Y\text{-}465\text{-}473}$ mutant peptides identified in the tumor of patient 2. All CTL lines responded strongly to target cells pulsed with relevant mutant peptides, whereas responses to target cells pulsed with corresponding wild-type peptides were generally low or negligible (fig. S5C, top row). Viable tumor material from patient 2 was scarce,

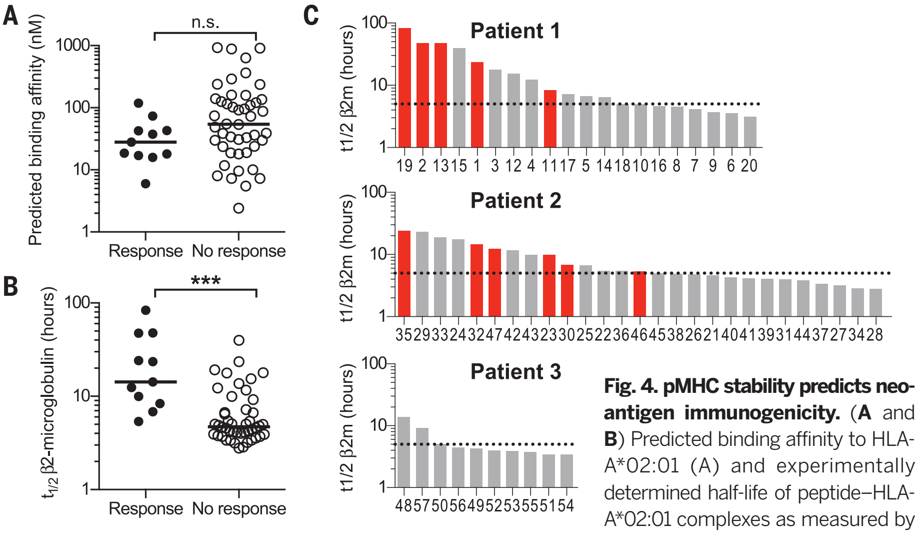


Fig. 4. pMHC stability predicts neoantigen immunogenicity. (**A** and **B**) Predicted binding affinity to HLA-A*02:01 (A) and experimentally determined half-life of peptide–HLA-A*02:01 complexes as measured by dissociation of β_2-microglobulin (B) for the 57 predicted neoantigens from patients 1, 2, and 3 that do or do not induce a T cell response. Peptide sequences and predicted affinities are listed in tables S1 to S3. (**C**) Red bars represent predicted neoantigens that were shown to be immunogenic; gray bars represent predicted neoantigens for which no T cell response could be detected. Dotted line represents suggested cutoff value of $t_{1/2}$ = 5 hours. Values in (B) and (C) represent means of triplicates. ***P < 0.0001 (Mann-Whitney U test), n.s., not significant.

and a tumor cell line for use in functional analyses could not be established. However, all but one of the CTL lines specifically recognized target cells transfected with a minigene encoding the mutant peptides PGM5$_{H>Y}$, USP28$_{C>F}$, and SNX24$_{P>L}$, flanked on both sides by 10 naturally occurring amino acids (fig. S5C, bottom row). Together, these data demonstrate that neoantigen-specific T cell responses can readily be induced in T cell repertoires from healthy donors, and that these T cells specifically recognize naturally processed neoantigens, including antigens expressed in matched tumor material.

Next, we investigated the feasibility of transferring donor-derived tumor-specific T cell reactivity by TCR gene transfer. TCRs from 28 CTL clones from three donors selected on the basis of reactivity toward melanoma cells of patient 1 were sequenced (table S4), yielding 11 unique TCR sequences, with one or two TCR sequences identified per antigen-donor combination. Nine of these were reconstructed and eight were successfully expressed in peripheral blood T cells, as confirmed by anti-TCRβ staining (fig. S6), targeting all four epitopes for which antitumor reactivity was seen. TCR-transduced PBMCs were then tested for degranulation and interferon (IFN)–γ production in response to cognate melanoma cells, with results for the five most responsive TCRs shown in Fig. 3.

The MLL-2$_{L>H}$–reactive TCR 41 strongly recognized patient-derived melanoma cells carrying the mutant MLL2 gene, with low reactivity to a third-party melanoma line lacking this mutation, unless the mutant epitope was genetically introduced (Fig. 3A). Furthermore, when the mutant MLL2 open reading frame in the cognate melanoma was disrupted, recognition of the mutant tumor was comparable to that of the third-party tumor (Fig. 3A). In addition, three CDK4$_{R>L}$-reactive TCRs were expressed in healthy donor T cells (TCR 53, 55, and 57), with two of these showing high recognition of the cognate melanoma that carries the mutant CDK4 gene as well as recognition of melanoma cell line Mel 526, which carries the previously described CDK4 R24C mutation (*17*) (Fig. 3B). Notably, recognition by TCRs 53 and 57 was comparable to that seen for the patient-derived CDK4$_{R>L}$-reactive TCR 17, previously isolated from TILs of patient 1 (Fig. 3, B and C, and fig. S7). The ASTN1$_{P>L}$-reactive TCR 65 also showed specific recognition of cognate melanoma (Fig. 3D). ASTN1$_{P>L}$-reactive TCR 52 and SMARCD3$_{H>Y}$-reactive TCRs 59 and 67 did not recognize cognate or third-party tumor unless the relevant neoantigen was introduced. In total, recognition of endogenously presented neoantigen on the cognate melanoma was observed for three of four antigens evaluated.

With neoantigens emerging as attractive targets in the development of personalized immunotherapies, strategies for the rapid identification of relevant neoantigens have become a major priority. We speculated that the use of outsourced immune responses could facilitate analysis of the rules that govern neoantigen recognition by T cells. In the current experiments, immunogenicity was evaluated for 57 peptides that had been selected on the basis of predicted binding affinity to HLA-A*02:01. Of these, 11 generated immune responses, and T cells reactive with 10 of these epitopes recognized endogenously presented antigen. The median predicted binding affinity for this set of T cell–recognized neoantigens was 28 nM (range 6 to 119 nM), compared to 54 nM (range 2 to 925 nM) for peptides that did not induce immune responses (Fig. 4A). Prior work has suggested that pMHC complex stability may form a particularly strong determinant of immunogenicity (*18*, *19*). To test the added value of experimental analysis of pMHC off-rate, we developed a flow cytometry–based assay for pMHC stability (fig. S8, A and B, and tables S1, S2, S3, and S5). Analysis of pMHC off-rates for all 57 predicted neoantigens revealed that neopeptides that were recognized by donor-derived T cells displayed a significantly longer half-life relative to neopeptides for which no responses were observed (median $t_{1/2}$ of β_2-microglobulin signal: 14.3 versus 4.7 hours, P < 0.0001) (Fig. 4B). Using a $t_{1/2}$ cutoff value of 5 hours, 11 of 32 (34%) candidate neoantigens were recognized by donor T cells (Fig. 4C). Furthermore, the significant added value of measured pMHC off-rates, as compared to the sole in silico prediction of peptide affinity, is also apparent from receiver operating characteristic (ROC) curves (fig. S8C).

Our results show that T cell repertoires from healthy donors provide a rich source of T cells that specifically recognize neoantigens present on human tumors. Responses to 11 different epitopes were observed, and for the majority of evaluated epitopes, potent and specific recognition of tumor cells endogenously presenting the neoantigens was detected. We draw three main conclusions from this work. First, these results demonstrate the existence of a repertoire of neoantigens on human tumors to which the endogenous T cell pool has not mounted a measurable response in vivo, but that can be the target of T cells from an independent source. Specifically, among the neoantigen-specific T cell populations capable of recognizing endogenously processed antigen, only one was also detected within the original TILs. This observation forms a strong incentive for the further development of immunotherapies that aim to broaden neoantigen-specific T cell reactivity (*14*, *20*, *21*), either from an exogenous source or from the endogenous T cell pool. (Note that the latter approach relies on the presence of patient T cells that still have the capacity to respond to these neglected epitopes, an issue that remains to be addressed.) Second, the ability to evaluate large series of predicted epitopes for recognition by T cells from multiple independent T cell repertoires makes it feasible to systematically examine the rules that control neoantigen recognition. Finally, the current results suggest the possibility of personalized neoantigen-directed immunotherapies that are independent of the status of the patient's own immune system.

REFERENCES AND NOTES

1. S. A. Rosenberg, N. P. Restifo, *Science* **348**, 62–68 (2015).
2. P. Sharma, J. P. Allison, *Science* **348**, 56–61 (2015).
3. N. A. Rizvi *et al.*, *Science* **348**, 124–128 (2015).
4. E. M. Van Allen *et al.*, *Science* **350**, 207–211 (2015).
5. P. F. Robbins *et al.*, *Nat. Med.* **19**, 747–752 (2013).
6. N. van Rooij *et al.*, *J. Clin. Oncol.* **31**, e439–e442 (2013).
7. C. Linnemann *et al.*, *Nat. Med.* **21**, 81–85 (2015).
8. E. Tran *et al.*, *Science* **350**, 1387–1390 (2015).
9. E. Tran *et al.*, *Science* **344**, 641–645 (2014).
10. M. M. Gubin *et al.*, *Nature* **515**, 577–581 (2014).
11. H. Matsushita *et al.*, *Nature* **482**, 400–404 (2012).
12. M. Yadav *et al.*, *Nature* **515**, 572–576 (2014).
13. D. Mittal, M. M. Gubin, R. D. Schreiber, M. J. Smyth, *Curr. Opin. Immunol.* **27**, 16–25 (2014).
14. B. M. Carreno *et al.*, *Science* **348**, 803–808 (2015).
15. See supplementary materials on *Science* Online.
16. I. W. Abrahamsen *et al.*, *Leukemia* **24**, 1901–1909 (2010).

17. T. Wölfel et al., Science 269, 1281–1284 (1995).
18. S. H. van der Burg, M. J. Visseren, R. M. Brandt, W. M. Kast, C. J. Melief, J. Immunol. 156, 3308–3314 (1996).
19. M. Harndahl et al., Eur. J. Immunol. 42, 1405–1416 (2012).
20. S. Kreiter et al., Nature 520, 692–696 (2015).
21. C. A. Klebanoff, S. A. Rosenberg, N. P. Restifo, Nat. Med. 22, 26–36 (2016).

ACKNOWLEDGMENTS

We thank L. Fanchi, R. Mezzadra, and F. Scheeren for advice and support; A. Vefferstad and the OUH flow cytometry core facility for exemplary technical assistance; I. M. Svane for sharing clinical material; and the Norwegian Bone Marrow Donor Registry for HLA typing. The data presented in the manuscript are tabulated in the main paper and in the supplementary materials. DNA and RNA sequencing data have been deposited in the European Genome-Phenome Archive (accession codes EGAD00001000243 and EGAD00001000325). A patent application (P32649NL00) has been filed that covers the technology of targeting cancer-specific amino acid sequences with donor-derived TCR repertoires (inventors J.O., T.N.S., E.S., and M.T.). All described TCRs are available under an MTA with the Netherlands Cancer Institute and Oslo University Hospital. Supported by Dutch Cancer Society Queen Wilhelmina Award NKI 2013-6122 and EU H2020 project APERIM (T.N.S.); the K. G. Jebsen Foundation (J.O., T.N.S., F.L.-J.); and the Research Council of Norway, Regional Authorities South-Eastern Norway, the University of Oslo and Oslo University Hospital, and the Norwegian Cancer Society (J.O.).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1337/suppl/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S8
References (22–43)

17 January 2016; accepted 4 May 2016
Published online 19 May 2016
10.1126/science.aaf2288

DEVELOPMENT

Zebrafish models of idiopathic scoliosis link cerebrospinal fluid flow defects to spine curvature

D. T. Grimes,[1]* C. W. Boswell,[2,3]* N. F. C. Morante,[1]* R. M. Henkelman,[4,5] R. D. Burdine,[1] B. Ciruna[2,3]†

Idiopathic scoliosis (IS) affects 3% of children worldwide, yet the mechanisms underlying this spinal deformity remain unknown. Here we show that *ptk7* mutant zebrafish, a faithful developmental model of IS, exhibit defects in ependymal cell cilia development and cerebrospinal fluid (CSF) flow. Transgenic reintroduction of Ptk7 in motile ciliated lineages prevents scoliosis in *ptk7* mutants, and mutation of multiple independent cilia motility genes yields IS phenotypes. We define a finite developmental window for motile cilia in zebrafish spine morphogenesis. Notably, restoration of cilia motility after the onset of scoliosis blocks spinal curve progression. Together, our results indicate a critical role for cilia-driven CSF flow in spine development, implicate irregularities in CSF flow as an underlying biological cause of IS, and suggest that noninvasive therapeutic intervention may prevent severe scoliosis.

Idiopathic scoliosis (IS) is a complex genetic disorder characterized by three-dimensional spinal curvatures, which arise in the absence of observable physiological or anatomical defects. Commonly diagnosed during adolescence, IS can cause disfigurement, reduced respiratory and pulmonary function, and chronic pain (*1*). In congenital and neuromuscular forms of scoliosis, spinal curves develop from vertebral malformations and/or underlying morbidities of the musculature and nervous system; however, the biological cause of IS has thus far remained unknown. As a result, treatment is limited to managing spinal deformity post-onset, through bracing and/or corrective surgery (*1*).

Genome-wide association studies have identified IS-associated polymorphisms in divergent human populations, but phenotypic and genetic variability have made it difficult to define causative mutations (*1*). Furthermore, a historical lack of appropriate animal models has confounded our basic understanding of the biology underlying IS (*2*). However, teleosts (bony fish) are, like humans, naturally prone to idiopathic spinal curvature (*3*), and recent genetic studies have identified faithful zebrafish IS models, providing important insights into the genetic causes of scoliosis (*4*, *5*) as well as a means to functionally validate human IS-associated genetic variants (*4*, *6*, *7*). Notably, zebrafish *ptk7* (protein tyrosine kinase-7) mutants present all defining attributes of the human disease, and studies of these mutants have implicated dysregulated Wnt signaling in the pathogenesis of IS (*4*).

Ptk7 is an essential regulator of both canonical Wnt–ß-catenin and noncanonical Wnt–planar cell polarity (PCP) signaling pathways (*8*). Although defects in either pathway are associated with a range of developmental abnormalities, both Wnt-PCP and Wnt–ß-catenin signaling have been implicated in the function of cilia (*9*–*11*). Cilia are microtubule-based organelles that project into the extracellular space and play critical roles in the perception and integration of environmental signals (*12*, *13*). Although most cell types elaborate short primary cilia, longer motile cilia are present on the surface of specialized cells and generate directional extracellular fluid flow in several contexts.

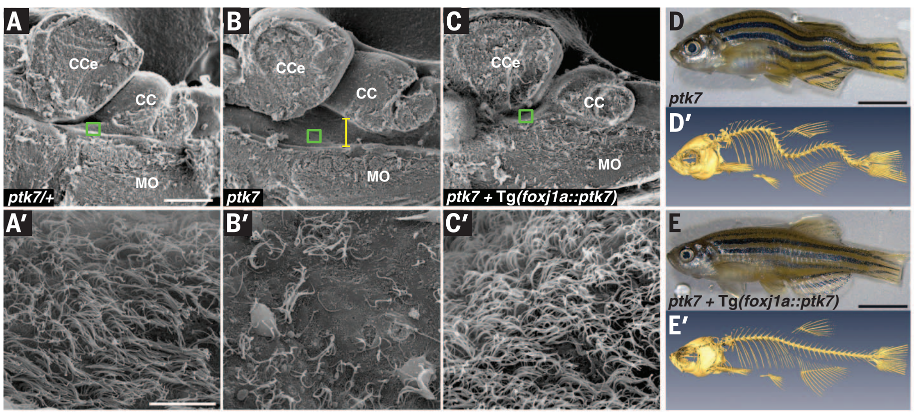


Fig. 1. *ptk7* mutant fish exhibit hydrocephalus, EC cilia defects, and spinal curves, all of which are prevented by transgenic reintroduction of *ptk7* specifically in motile ciliated cell lineages. (**A** to **C**) Representative sagittal SEM images of the brains of a *ptk7*/+ control [*n* = 6 (A)], a *ptk7* mutant [*n* = 6 (B)], and a *ptk7* mutant expressing Tg(*foxj1a*::*ptk7*) [*n* = 6 (C)], all at 2.5 months of age. The yellow line in (B) demarcates hydrocephalus. Green squares indicate the areas shown in corresponding high-magnification SEM images (**A′** to **C′**). (**D** to **E′**) Representative fixed [(D) and (E)] and μCT-rendered [(D′) and (E′)] lateral views of an adult *ptk7* mutant [(D) and (D′)] and an adult *ptk7* mutant expressing Tg(*foxj1a*::*ptk7*) [(E) and (E′)]. CCe, corpus cerebelli; CC, crista cerebellaris; MO, medulla oblongata. Scale bars, 250 μm [(A), (B), and (C)], 10 μm [(A′), (B′), and (C′)], and 5 mm [(D) and (E)].

[1]Department of Molecular Biology, Princeton University, Washington Road, Princeton, NJ 08544, USA. [2]Program in Developmental & Stem Cell Biology, The Hospital for Sick Children, 686 Bay Street, Toronto, Ontario M5G 0A4, Canada. [3]Department of Molecular Genetics, The University of Toronto, Toronto, Ontario M5S 1A8, Canada. [4]Mouse Imaging Centre, The Hospital for Sick Children, 25 Orde Street, Toronto, Ontario M5T 3H7, Canada. [5]Department of Medical Biophysics, The University of Toronto, Toronto, Ontario M5G 2M9, Canada.
*These authors contributed equally to this work. **†Corresponding author. Email: ciruna@sickkids.ca**

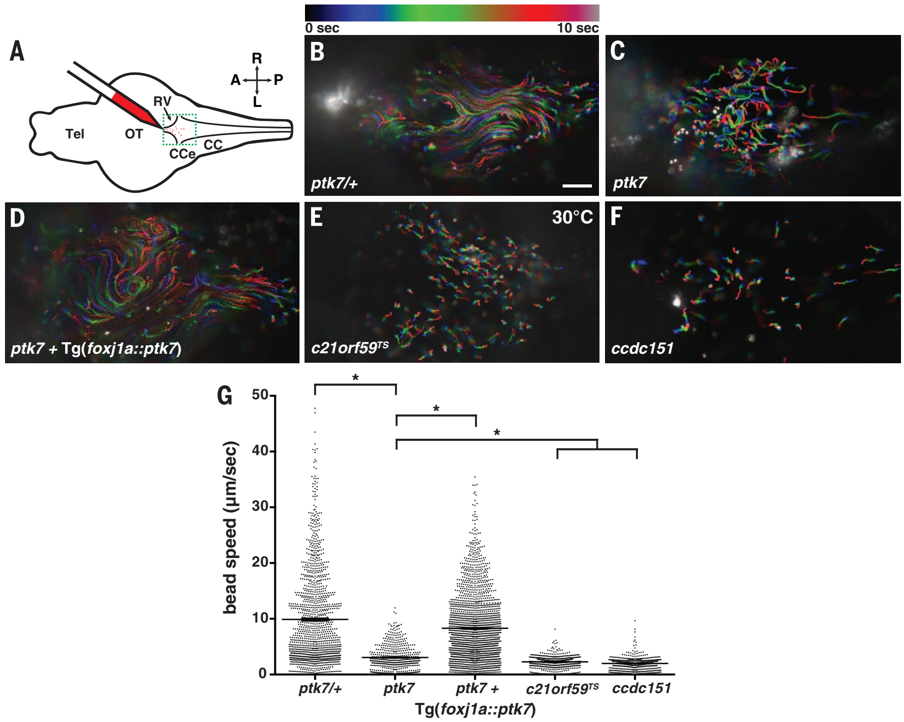


Fig. 2. CSF flow is compromised in zebrafish IS models. (**A**) Schematic of ventricular flow assay in whole-mount adult brains. The green dashed box represents the area imaged. (**B** and **C**) Bulk trajectory patterns of beads in whole-mount adult brains of *ptk7/+* controls (B) and *ptk7* mutants (C). (**D**) Trajectory pattern of *ptk7* mutants expressing Tg(*foxj1a::ptk7*). (**E** and **F**) Trajectory patterns in additional IS models. In (B) to (F), trajectory paths are color-coded to represent initial position (blue) and final position (red) over 10 s. (**G**) Quantification of CSF flow in various IS models and controls, with data points representing an individual bead speed. Bars represent means. Standard errors of the means are as follows: *ptk7/+*, ±0.2352; *ptk7*, ±0.1013; *ptk7*+ Tg (*foxj1a::ptk7*), ± 0.1363; *c21orf59^{TS}*, ± 0.0374; and *ccdc15*, ±0.0342. Comparisons between genotypes used a *t* test. **P* < 0.0001. A, anterior; L, left; OT, optic tectum; P, posterior; RV, rhombencephalic ventricle; R, right; Tel, telencephalon. Scale bar, 50 μm [(B) to (F)].

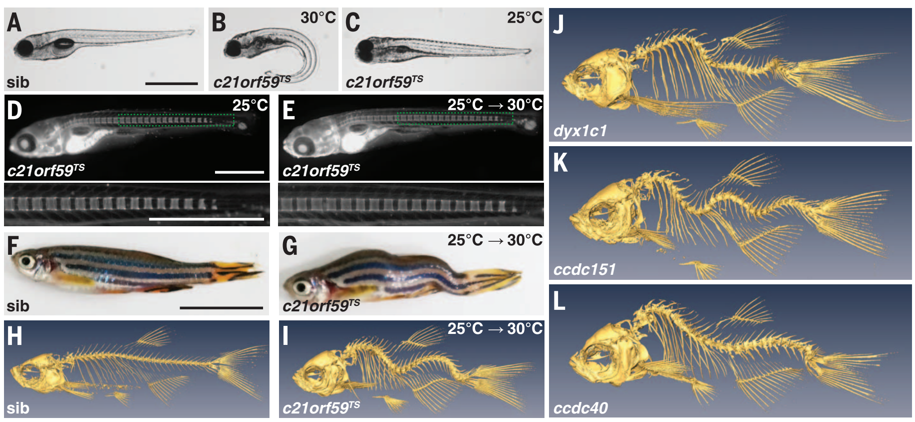


Fig. 3. Cilia motility mutants exhibit spinal curves. (**A** to **C**) Whereas sibling (sib) controls had no phenotype (A), mutant larvae from *c21orf59^{TS}* intercrosses at 3 dpf exhibited cilia motility–associated defects, including ventral axis curvature, at 30°C (B) but not at 25°C (C). (**D** and **E**) Calcein staining of *c21orf59^{TS}* mutants repaired during embryogenesis revealed no vertebral malformations during larval stages at either permissive [*n* = 4 (D)] or restrictive [*n* = 7 (E)] temperatures. Green rectangles indicate the areas shown in higher magnification below. (**F** and **G**) *c21orf59^{TS}* mutants raised at 25°C until 5 dpf and then shifted to 30°C exhibited spinal curves at sexual maturity [*n* = 10 (G)] that were not present in sib controls [*n* = 3 (F)]. (**H** and **I**) μCT of sib controls [*n* = 3 (H)] and *c21orf59^{TS}* mutants raised at 25°C until 5 dpf and then shifted to 30°C [*n* = 10 (I)]. (**J** to **L**) μCT of *dyx1c1* mutants [(*n* = 3 (J)], *ccdc151* mutants [*n* = 7 (K)], and *ccdc40* mutants [*n* = 3 (L)], repaired during embryonic stages by injection of wild-type mRNA at the one-cell stage. Scale bars, 1 mm [(A) to (E)] and 1 cm [(F) and (G)].

Cilia-directed flow within early embryonic organizers breaks left-right (L-R) symmetry in development (*14*), and cerebrospinal fluid (CSF) flow, which is critical for central nervous system homeostasis (*15*), is generated by the polarized beating of ependymal cell (EC) cilia lining brain ventricles (*16*). Abnormal L-R asymmetries and defective CSF flow have been observed in IS patients (*17*), and an elevated incidence of scoliosis has been documented among primary ciliary dyskinesia patients (*18*). We therefore hypothesized that motile cilia dysfunction may contribute to the etiopathogenesis of IS.

To test this, we first investigated EC motile cilia structure and function in scoliotic *ptk7* mutant zebrafish and sibling *ptk7*/+ controls. Examination of *ptk7* mutant brain ventricles by scanning electron microscopy (SEM) revealed severe hydrocephalus (Fig. 1, A and B), a phenotype commonly associated with loss of EC cilia function (*16*). Moreover, whereas a dense network of polarized EC cilia lined the ventral surface of *ptk7*/+ ventricles, cilia in *ptk7* mutant ventricles were sparse and, when present, lacked posterior polarization (Fig. 1, A′ and B′). To directly examine the consequence of EC cilia defects, we tracked fluorescent microsphere movement across the ventral surface of the rhombencephalic ventricle (Fig. 2A). Dynamic anterior-to-posterior flow was observed across the ventricle of *ptk7*/+ brains (Fig. 2, B and G, and movie S1). In contrast, although some movement was observed during particle tracking in *ptk7* mutants, microspheres exhibited irregular trajectories and significantly reduced speeds (Fig. 2, C and G, and movie S2). These results demonstrate abnormal CSF flow within the ventricular system of scoliotic *ptk7* mutants and are consistent with a role for EC motile cilia defects in the etiology of IS.

To investigate whether scoliosis specifically results from motile cilia dysfunction, we assessed potential amelioration of *ptk7* mutant spinal curves through transgenic reintroduction of wild-type Ptk7 in motile ciliated cell lineages only. The transcription factor FoxJ1a is a master regulator of motile ciliogenesis (*19*). We therefore cloned and characterized a *foxj1a* enhancer element that specifically drives transgene (Tg) expression in all known sites of motile cilia formation, as demonstrated in multiple *foxj1a::eGFP* transgenic lines (eGFP, enhanced green fluorescent protein) (fig. S1, A to D, and movie S3). Along the trunk of juvenile animals, Tg(*foxj1a::eGFP*) expression was predominantly restricted to midline structures of the brain and spinal cord. We next generated four independent *foxj1a::ptk7* stable transgenic lines (fig. S1E) and found that the presence of Tg(*foxj1a::ptk7*) restored EC cilia and CSF flow in *ptk7* mutant fish and prevented hydrocephalus from manifesting (Fig. 1, C and C′; Fig. 2, D and G; and movie S4). Importantly, spinal curve formation, assessed by microcomputed tomography (μCT), was also fully suppressed by the transgenes (*n* = 59; Fig. 1, D to E′), showing that scoliosis in mutants is specifically caused by Ptk7 dysfunction in motile ciliated lineages.

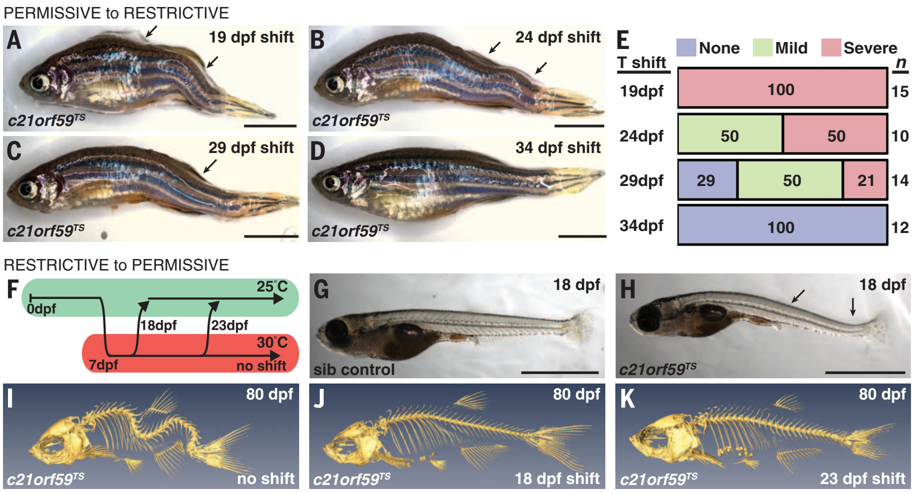


Fig. 4. Temporal window for motile cilia function in spine development. (**A** to **D**) Representative lateral views of *c21orf59*^{TS} mutants that were shifted from 25°C to 30°C at 19 dpf (A), 24 dpf (B), 29 dpf (C), and 34 dpf (D). (**E**) Quantification of scoliosis phenotypes in temperature-shifted mutants, observed at 6 months post-fertilization (numbers are percentages). (**F**) Schematic of restrictive (30°C)–to–permissive (25°C) temperature shift experiments performed with *c21orf59*^{TS} mutants. (**G** and **H**) Lateral images of a juvenile sib control (G) and a *c21orf59*^{TS} mutant (H), both kept at 30°C, show curve initiation in mutants by 18 dpf. (**I**) Representative μCT image of an 80-dpf unshifted *c21orf59*^{TS} mutant. (**J** and **K**) Representative μCT images of 80-dpf *c21orf59*^{TS} mutants shifted from 30°C to 25°C at 18 dpf [n = 6 (J)] or at 23 dpf [n = 6 (K)]. Scale bars, 5 mm [(A) to (D)] and 2 mm [(G) and (H)].

If cilia motility defects contribute to IS pathogenesis, then *ccdc40* (*20*), *ccdc151* (*21*), *dyx1c1* (*22*), and *c21orf59* (*23*, *24*) mutations, which all disrupt cilia motility, should lead to the development of scoliosis. However, aberrant cilia motility causes a characteristic suite of embryonic phenotypes that usually result in death by 1 to 2 weeks of development (*20*, *21*), precluding analysis of adolescent spine formation. To circumvent this early lethality, we used two strategies. First, we took advantage of the *c21orf59* temperature-sensitive mutation *tm304*, here called *c21orf59*^{TS} (*24*). At 30°C (a restrictive temperature), *c21orf59*^{TS} mutant embryos exhibited abnormal cilia motility and associated developmental defects (Fig. 3, A and B). However, at 25°C (a permissive temperature), *c21orf59*^{TS} embryos retained cilia motility and could develop normally (Fig. 3C) (*24*). *c21orf59*^{TS} mutants that were raised at 25°C for 5 days to prevent embryonic defects and then shifted to 30°C resembled wild-type zebrafish through the larval stages, exhibiting normal vertebral formation, as monitored using the vital fluorescent Ca^{2+}-binding chromophore calcein (Fig. 3, D and E). CSF flow in the rhombencephalic ventricle of these *c21orf59*^{TS} mutants was severely compromised (Fig. 2, E and G, and movie S5). Moreover, all mutant fish developed spinal curves that began to form during early juvenile stages (3 to 4 weeks of age; fig. S2, A and B) and that model defining attributes of IS (Fig. 3, F to I, and movie S6).

Our second strategy involved suppressing embryonic phenotypes by means of RNA injections at the one-cell stage and analyzing mutants during adolescence. Using CRISPR/Cas9 gene targeting, we generated a *dyx1c1* mutant allele (fig. S3). Functional characterization of *dyx1c1* mutants revealed abnormal cilia motility and associated developmental defects, including embryonic lethality (fig. S3, C to F), in agreement with gene knockdown studies (*22*). *dyx1c1* mutants injected with wild-type mRNA to prevent embryonic defects developed severe three-dimensional spinal curvatures in the absence of congenital vertebral malformations (Fig. 3J; fig. S2, A and B; and fig. S4). Furthermore, *ccdc151* and *ccdc40* mutant embryos that were phenotypically normal in embryonic stages (owing to wild-type mRNA injection) also developed late-onset spinal curves that model IS (Fig. 3, K to L). Our demonstration that mutations in four different genes, each of which has been shown to disrupt cilia motility, all yield similar adolescent spinal curve phenotypes provides strong evidence that motile cilia dysfunction represents the underlying cell-biological cause of IS in these models.

These experiments further demonstrate a post-embryonic requirement for motile cilia in spine morphogenesis. Transient knockdown of Dyx1c1 or Ccdc151 through only the first 3 to 4 days of embryogenesis [by injection of translation-blocking antisense morpholino oligonucleotides (MOs)] did not result in adolescent spinal curvatures, despite the fact that MO-injected embryos phenocopied genetic mutants during early embryogenesis (fig. S5). To define the critical developmental window for motile cilia function in the etiopathogenesis of IS, we performed a series of temperature shift experiments using the *c21orf59*^{TS} mutant allele. *c21orf59*^{TS} mutant embryos were raised at 25°C for at least 5 days (to prevent embryonic phenotypes), transferred to a restrictive temperature (30°C) at defined incremental stages of development, and screened for spinal curvatures at sexual maturity (Fig. 4, A to D). *c21orf59*^{TS} mutants that were shifted to restrictive temperatures at 19 days post-fertilization (dpf) all developed severe spinal curves by 5 weeks of age (Fig. 4, A and E). In contrast, *c21orf59*^{TS} mutants that were shifted to 30°C at 24 and 29 dpf exhibited milder spinal curvatures (Fig. 4, B, C, and E), whereas *c21orf59*^{TS} mutants that were shifted to 30°C at 34 dpf displayed no signs of scoliosis through the adult stages (Fig. 4, D and E). These results indicate a finite and temporally defined requirement for motile cilia function during spine morphogenesis. This time interval correlates with documented periods of accelerated adolescent growth (*4*), when spinal curves typically manifest in IS.

Last, to determine whether restoration of motile cilia function can prevent severe spinal curve progression after the onset of scoliosis, we performed restrictive-to-permissive temperature shifts at defined time points. *c21orf59*^{TS} mutant embryos were first raised at 25°C until 7 dpf to allow normal embryonic development, transferred to 30°C until the onset of spinal curve formation, and then returned to permissive temperatures at incremental stages of spinal curve progression (Fig. 4, F to H). Restoration of motile cilia activity at the onset of scoliosis blocked spinal curve progression (Fig. 4, J to K). This provides a proof-of-principle that the development of severe IS spinal curvatures can be managed without invasive surgical manipulation.

The data presented here demonstrate that cilia motility is required for zebrafish spine morphogenesis. Given the acute hydrocephalus and EC cilia defects observed in *ptk7* mutants, the predominant expression of *foxj1a* transgenes throughout the brain and spinal cord of juvenile animals, and the severe CSF flow defects observed across zebrafish IS models, we suggest that irregularities in CSF flow represent the underlying cell-biological cause of IS. Several observations support this model: (i) Disruption of CSF activity via Kaolin injection into the subarachnoid space can cause scoliosis in both dog and rabbit models (*25*, *26*), and (ii) scoliosis is highly prevalent in multiple human conditions associated with obstructed CSF flow, including Chiari malformation, syringomyelia, and myelomeningoceles (*27*–*29*). Our data explain these observations and further imply an evolutionarily conserved role for CSF flow in spine morphogenesis, thus warranting reexamination of the anatomy, physiology, and genetics of CSF flow in cases of human IS. Downstream of CSF flow, molecular mechanisms influencing spine morphogenesis remain to be determined but could involve multiple gene products that have been previously associated with IS [e.g., potential motile cilia functions for the centriolar protein POC5 (*6*) or chondrocyte-specific activation of GPR126 (*30*)]. Ultimately, our demonstration that severe spinal curvatures can be prevented with the restoration of motile cilia activity may have important therapeutic ramifications; pharmaceutical manipulation of the production and/or downstream interpretation of CSF signals could potentially stop severe

spinal curve progression in some IS patients, even after the onset and clinical diagnosis of scoliosis.

REFERENCES AND NOTES

1. J. C. Cheng et al., Nat. Rev. Dis. Prim. 1, 15030 (2015).
2. M. M. Janssen, R. F. de Wilde, J. W. Kouwenhoven, R. M. Castelein, Spine J. 11, 347–358 (2011).
3. K. F. Gorman, F. Breden, Med. Hypotheses 72, 348–352 (2009).
4. M. Hayes et al., Nat. Commun. 5, 4777 (2014).
5. J. G. Buchan et al., Dev. Dyn. 243, 1646–1657 (2014).
6. S. A. Patten et al., J. Clin. Invest. 125, 1124–1128 (2015).
7. S. Sharma et al., Nat. Commun. 6, 6452 (2015).
8. M. Hayes, M. Naito, A. Daulat, S. Angers, B. Ciruna, Development 140, 1807–1818 (2013).
9. T. J. Park, B. J. Mitchell, P. B. Abitua, C. Kintner, J. B. Wallingford, Nat. Genet. 40, 871–879 (2008).
10. A. Borovina, S. Superina, D. Voskas, B. Ciruna, Nat. Cell Biol. 12, 407–412 (2010).
11. A. Caron, X. Xu, X. Lin, Development 139, 514–524 (2012).
12. V. Singla, J. F. Reiter, Science 313, 629–633 (2006).
13. S. C. Goetz, K. V. Anderson, Nat. Rev. Genet. 11, 331–344 (2010).
14. J. J. Essner et al., Nature 418, 37–38 (2002).
15. M. J. Simon, J. J. Iliff, Biochim. Biophys. Acta 1862, 442–451 (2016).
16. L. Lee, J. Neurosci. Res. 91, 1117–1132 (2013).
17. W. J. Wang et al., J. Pediatr. Orthop. 31, S14–S27 (2011).
18. V. G. Engesaeth, J. O. Warner, A. Bush, Pediatr. Pulmonol. 16, 9–12 (1993).
19. S. P. Choksi, G. Lauter, P. Swoboda, S. Roy, Development 141, 1427–1441 (2014).
20. A. Becker-Heck et al., Nat. Genet. 43, 79–84 (2011).
21. R. Hjeij et al., Am. J. Hum. Genet. 95, 257–274 (2014).
22. A. Tarkar et al., Nat. Genet. 45, 995–1003 (2013).
23. C. Austin-Tse et al., Am. J. Hum. Genet. 93, 672–686 (2013).
24. K. M. Jaffe et al., Cell Rep. 14, 1841–1849 (2016).
25. A. Chuma et al., Spine 22, 589–594, discussion 595 (1997).
26. M. Turgut, E. Cullu, A. Uysal, M. E. Yurtseven, B. Alparslan, Neurosurg. Rev. 28, 289–297 (2005).
27. T. H. Milhorat et al., Neurosurgery 44, 1005–1017 (1999).
28. R. A. Ozerdemoglu, F. Denis, E. E. Transfeldt, Spine 28, 1410–1417 (2003).
29. M. Verhoef et al., Dev. Med. Child Neurol. 46, 420–427 (2004).
30. C. M. Karner, F. Long, L. Solnica-Krezel, K. R. Monk, R. S. Gray, Hum. Mol. Genet. 24, 4365–4373 (2015).

ACKNOWLEDGMENTS

We gratefully acknowledge E. Lee, D. Holmyard, P. Paroutis, and L. Yu for technical assistance; A. Morley, A. Ng, C. Hasty, D. Bosco, and P. Johnson for zebrafish care; and J. Schottenfeld-Roames and T. Ku for initial observations of spine curvatures in *c21orf59^{TS}* mutant fish. This work was supported in part by funding from the Canada Research Chairs program to R.M.H. and B.C.; National Institute of Child Health and Development grant 2R01HD048584 to R.D.B.; and Canadian Institutes of Health Research (MOP-111075) and March of Dimes Foundation (#1-FY13-398) grants to B.C. D.T.G., C.W.B., and N.F.C.M. performed all experiments characterizing scoliosis phenotypes. C.W.B. cloned and characterized *foxj1a* transgenic lines. N.F.C.M. generated the *dyx1c1^{pr13}* allele. C.W.B. and R.M.H. performed μCT. B.C. and C.W.B. performed SEM and CSF flow analyses. D.T.G., C.W.B., N.F.C.M., R.D.B., and B.C. analyzed the data. D.T.G., C.W.B. and B.C. wrote the manuscript, and all authors edited the manuscript.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1341/suppl/DC1
Materials and Methods
Figs. S1 to S5
Movies S1 to S6
References (31–33)

8 March 2016; accepted 16 May 2016
10.1126/science.aaf6419

CANCER

The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas

Dong Fang,[1]* Haiyun Gan,[1]* Jeong-Heon Lee,[1,2]* Jing Han,[1]* Zhiquan Wang,[1]* Scott M. Riester,[3] Long Jin,[3] Jianji Chen,[4] Hui Zhou,[1] Jinglong Wang,[5,6] Honglian Zhang,[1] Na Yang,[5] Elizabeth W. Bradley,[3] Thai H. Ho,[7] Brian P. Rubin,[8] Julia A. Bridge,[9] Stephen N. Thibodeau,[10] Tamas Ordog,[2,11,12] Yue Chen,[4] Andre J. van Wijnen,[1,3] Andre M. Oliveira,[3,10] Rui-Ming Xu,[5,6] Jennifer J. Westendorf,[1,3] Zhiguo Zhang[1,2]†

More than 90% of chondroblastomas contain a heterozygous mutation replacing lysine-36 with methionine-36 (K36M) in the histone H3 variant H3.3. Here we show that H3K36 methylation is reduced globally in human chondroblastomas and in chondrocytes harboring the same genetic mutation, due to inhibition of at least two H3K36 methyltransferases, MMSET and SETD2, by the H3.3K36M mutant proteins. Genes with altered expression as well as H3K36 di- and trimethylation in H3.3K36M cells are enriched in cancer pathways. In addition, H3.3K36M chondrocytes exhibit several hallmarks of cancer cells, including increased ability to form colonies, resistance to apoptosis, and defects in differentiation. Thus, H3.3K36M proteins reprogram the H3K36 methylation landscape and contribute to tumorigenesis, in part through altering the expression of cancer-associated genes.

Chondroblastomas are locally recurrent primary bone tumors (*1*). Recently, it has been reported that one allele of the *H3F3B* gene, one of two genes encoding histone H3 variant H3.3 (*2*, *3*), is frequently mutated in chondroblastoma (*4*). In addition, global reductions of di- and trimethylation of histone H3 at lysine-36 (H3K36me2 and H3K36me3) of endogenous histone H3 in mammalian cells exogenously expressing the H3.3 $Lys^{36} \rightarrow Met^{36}$ (H3.3K36M) mutant protein have been observed (*5*, *6*). However, the mechanism by which the mutant proteins exert their effects on H3K36 methylation of endogenous histones and how the H3.3K36M mutation promotes tumorigenesis of this poorly studied tumor are largely unknown.

We used H3K36me2- and H3K36me3-specific antibodies (fig. S1) to analyze the levels of H3K36me2 and H3K36me3 in three primary human chondroblastomas harboring the H3.3K36M mutation and in three giant cell tumors with the H3.3G34W ($Gly^{34} \rightarrow Trp^{34}$) mutation (table S1). H3K36me2 and H3K36me3 were globally reduced in each chondroblastoma specimen, but not in giant cell tumors or normal bone tissues (Fig. 1A).

We used the CRISPR/Cas9 system (*7*) to introduce the H3.3K36M mutation into one *H3F3B* allele (fig. S2) of T/C28a2 cells, which are immortalized human chondrocytes (*8*). The levels of H3K36me1, H3K36me2, and H3K36me3 were reduced in two independent mutant cell lines, as compared with methylation levels in parental T/C28a2 cells (Fig. 1B and fig. S3). As determined by mass spectrometric analysis, H3K36me2 was reduced more substantially than H3K36me3 (fig. S3, A and B). No apparent changes were observed in H3K4me3, H3K9me3, H3K27me3, or H4K20me3 (Fig. 1B). These results demonstrate that the global reduction of H3K36 methylation in tumor tissues results from the expression of H3.3K36M mutant proteins.

To understand how H3.3K36M mutant proteins globally reduce H3K36 methylation, we first tested the ability of a H3.3K36M peptide to inhibit the enzymatic activities of four human H3K36 methyltransferases—SETD2, ASH1L, MMSET/WHSC1, and NSD1—which catalyze H3K36me1, H3K36me2, and H3K36me3 (*9*–*11*). The purified catalytic domains of each enzyme exhibited methyltransferase activities against H3.3-containing mononucleosomes. The H3.3K36M peptide inhibited the

[1]Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA. [2]Epigenomics Program, Center of Individualized Medicine, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA. [3]Department of Orthopedic Surgery, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA. [4]Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota at Twin Cities, Minneapolis, MN 55455, USA. [5]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 5 Datun Road, Beijing 100101, China. [6]University of Chinese Academy of Sciences, 19A Yuquan Road, Beijing 100049, China. [7]Division of Hematology/Oncology, Mayo Clinic Arizona, 13400 East Shea B., Scottsdale, AZ 85259, USA. [8]Robert J. Tomsich Pathology and Laboratory Medicine Institute and Department of Cancer Biology, Cleveland Clinic and Lerner Research Institute, L2 9500 Euclid Avenue, Cleveland, OH 44195, USA. [9]Departments of Pathology and Microbiology, Pediatrics, and Orthopaedic Surgery and Rehabilitation. [10]Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA. [11]Department of Physiology and Biomedical Engineering, Division of Gastroenterology and Hepatology, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, USA. [12]Interdisciplinary Health Science Initiative, 1110 Micro and Nanotechnology Laboratory, M/C 249, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA.
*These authors contributed equally to this work. **†Corresponding author. Email: zhang.zhiguo@mayo.edu**

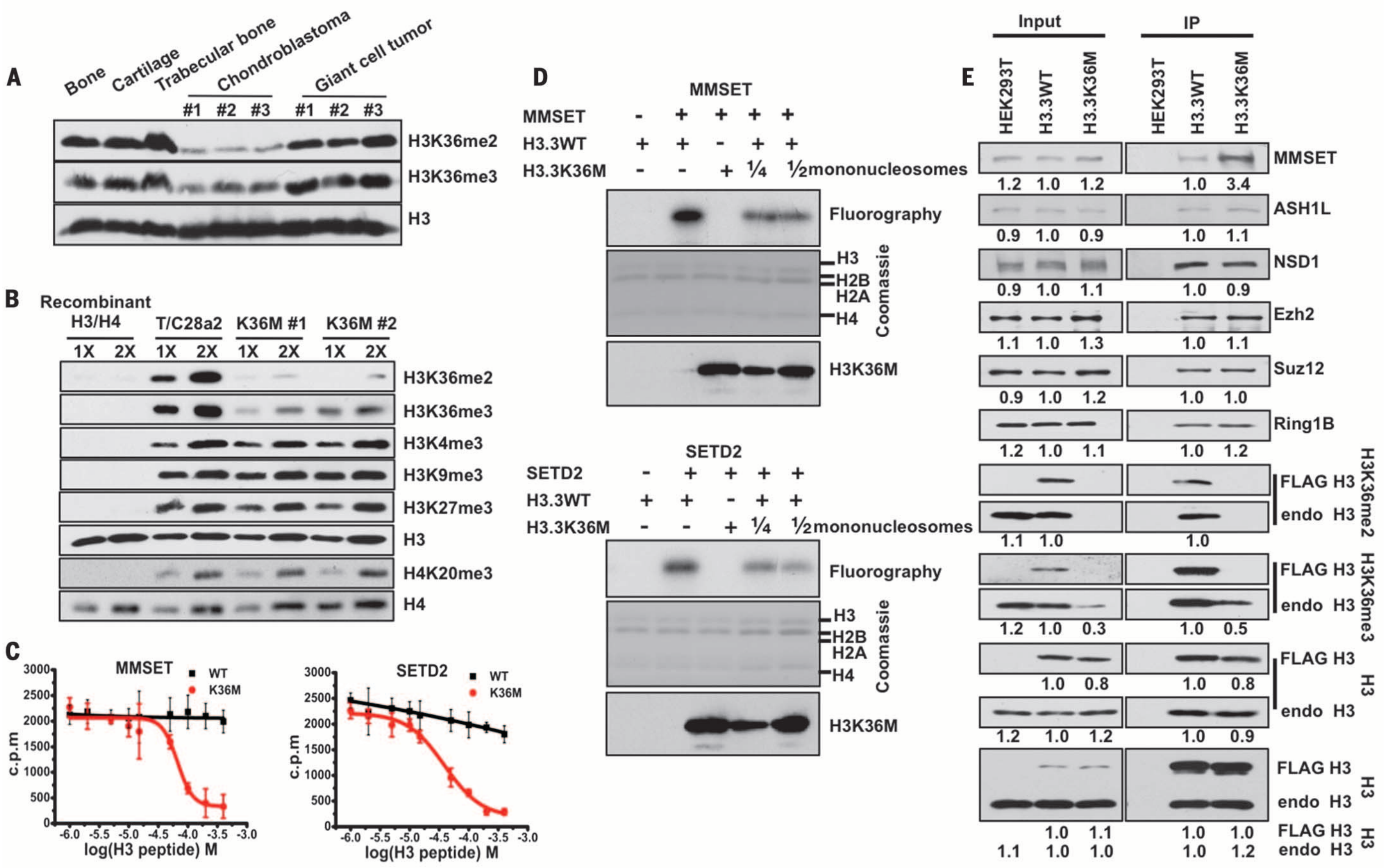


Fig. 1. H3K36 methylation is reduced in tumors and cells containing the H3.3K36M mutation. (**A** and **B**) H3K36me2 and H3K36me3 levels are reduced in chondroblastoma tumor samples (A) and in chondrocyte cell lines containing the H3.3K36M mutation (B). (**C**) Enzymatic activity of the H3K36 methyltransferases MMSET and SETD2 was measured using H3.3-containing recombinant mononucleosomes in the presence of increasing amounts of H3.3K36M peptide or its corresponding H3.3 peptide. Data are mean ± SD (N = 3 independent replicates). c.p.m., counts per minute. (**D**) H3.3K36M mononucleosomes inhibit the enzymatic activities of MMSET (top) and SETD2 (bottom) in vitro. (**E**) MMSET is enriched in H3.3K36M-containing mononucleosomes compared with H3.3 WT nucleosomes. Mononucleosomes were purified from human embryonic kidney 293 T (HEK293T) cells (negative control), HEK293T cells expressing FLAG-tagged WT H3.3, or the H3.3K36M mutant. Proteins from input and immunoprecipitated (IP) samples were analyzed by Western blotting using the indicated antibodies. The intensity of each blot was quantified, with blots of both input and IP in WT cells set as 1.0.

activities of MMSET [median inhibitory concentration (IC_{50}) = 67 μM] and SETD2 (IC_{50} = 39 μM) in a dose-dependent manner, as compared with the wild-type (WT) H3 peptide (Fig. 1C and fig. S4A). Moreover, H3.3K36M-containing mononucleosomes inhibited the enzymatic activities of MMSET and SETD2 (Fig. 1D). In contrast, neither the H3.3K36M peptide nor the H3.3K36M mononucleosomes exhibited an inhibitory effect on the activities of ASH1L and NSD1 in vitro (fig. S4). We also observed that MMSET—but not ASH1L, NSD1, or two subunits of the H3K27 methyltransferase complex PRC2 (Ezh2 and Suz12)—was enriched in H3.3K36M-containing mononucleosomes compared with WT H3.3-containing mononucleosomes (Fig. 1E). Finally, in a peptide pull-down assay, MMSET and SETD2 bound to the H3.3K36M peptide more efficiently under higher-salt conditions than to the corresponding normal H3 peptide (fig. S5A). These results indicate that the H3.3K36M mutant protein inhibits at least two mammalian H3K36 methyltransferases, MMSET and SETD2.

We used chromatin immunoprecipitation coupled with next-generation sequencing (ChIP-seq) (*12*) to identify 29,250 and 13,694 H3K36me2 peaks in T/C28a2 cells and H3.3K36M lines, respectively (Fig. 2, A and B). H3K36me2 peaks present in only the H3.3K36M cells were significantly enriched at promoters, gene bodies, and transcription end sites (TESs) ± 2 kilo–base pairs (kbp) (P < 0.01) but showed significant depletion at intergenic regions (P < 0.01), as compared with H3K36me2 peaks in T/C28a2 cells (Fig. 2, A to C). On average, levels of H3K36me2 in intergenic regions (Fig. 2D) and gene bodies (Fig. 2E) were reduced in each of the H3.3K36M mutant lines compared with the WT T/C28a2 cells. The reduction of H3K36me2 in four selected genes and two intergenic regions was confirmed by ChIP–polymerase chain reaction (PCR) (Fig. 2F).

Depletion of SETD2 had no apparent effect on H3K36me2. In contrast, depletion of MMSET alone or in combination with SETD2 resulted in a marked reduction of H3K36me2, but to a lesser extent than in cells expressing H3.3K36M, as determined from Western blot analysis (fig. S5, B and C). Depletion of SETD2 or MMSET did not affect the expression of the three other H3K36 methyltransferases that we tested (fig. S5D). Finally, H3K36me2 ChIP-PCR and ChIP-seq results show that depletion of MMSET alone or in combination with SETD2, but not SETD2 alone, led to reduction of H3K36me2 in gene bodies and intergenic regions but to a lesser extent than in H3.3K36M mutant cells (Fig. 2G and fig. S5, E and F). Together, these results support the idea that the reduction of H3K36me2 in H3.3K36M cells is mediated, at least in part, through the inhibition of MMSET by the H3.3K36M proteins.

Similar numbers of H3K36me3 ChIP-seq peaks were detected in gene bodies from T/C28a2 cells and H3.3K36M mutant cells (fig. S6A). However, the amount of H3K36me3 throughout the gene bodies was reduced in each of the two H3.3K36M mutant lines compared with the T/C28a2 cells (Fig. 2H). Additionally, more than 60% of genes with reduced levels of H3K36me3 also exhibited reduced H3K36me2 (fig. S6B). The reduction

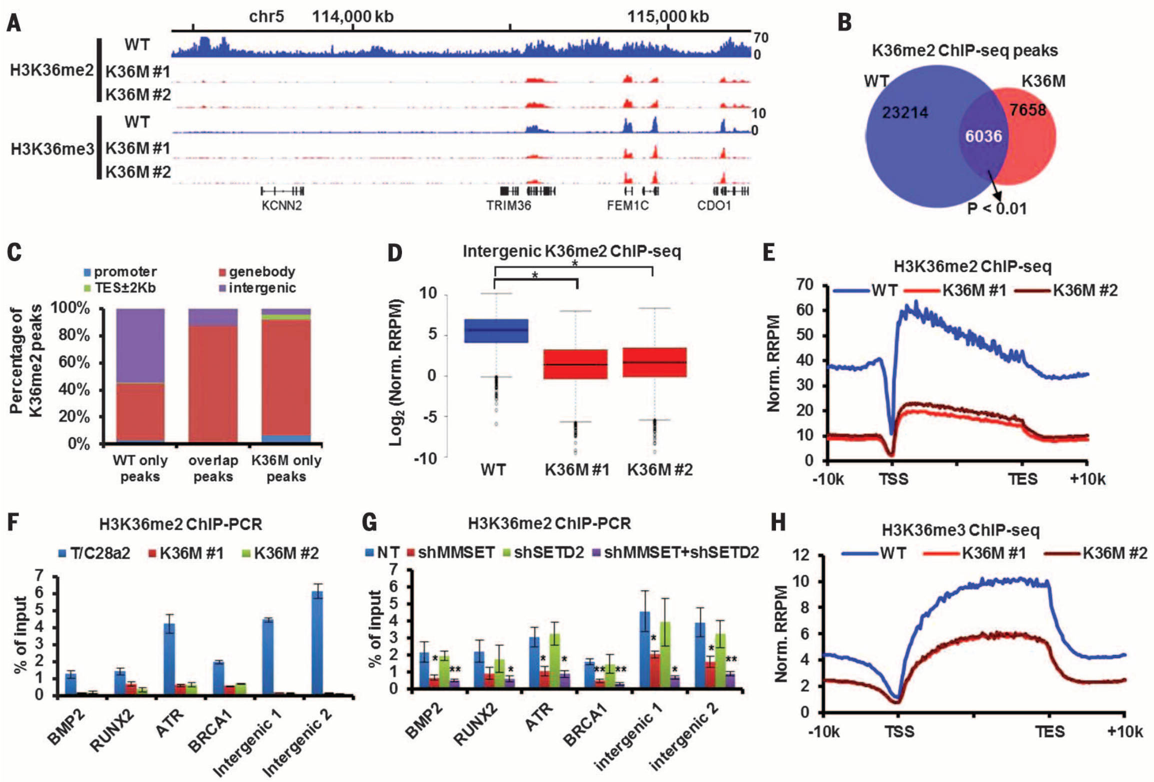


Fig. 2. H3.3K36M reprograms chromatin-bound H3K36me2 and H3K36me3. (**A**) Integrative Genomics Viewer tracks showing H3K36me2 and H3K36me3 distribution on chromosome 5 (chr5) in WT (blue) and two H3.3K36M (K36M) cell lines (red). (**B**) Venn diagram illustration of H3K36me2 peaks in WT and two H3.3K36M cell lines and their overlap. (**C**) Genomic distribution of H3K36me2 peaks unique to WT cells (23214), peaks shared by WT and H3.3K36M cells (6036), and peaks unique to H3.3K36M cells (7658). (**D** and **E**) Normalized tag distribution profiles of H3K36me2 for intergenic regions [(D), *$P <$ 0.01] and gene bodies (E) in WT and K36M cell lines. RRPM, reference-adjusted reads per million; TSS, transcription start site. (**F**) Analysis of H3K36me2 at four gene loci and two intergenic regions using ChIP-PCR in two H3.3K36M mutant lines. Error bars indicate SD from three independent replicates. (**G**) Depletion of MMSET, but not SETD2, reduced H3K36me2 at four gene loci and two intergenic regions. Data are mean ± SD (N = 3 independent replicates, *$P <$ 0.05, **$P <$ 0.01). NT, nontarget control. (**H**) H3K36me3 levels in gene bodies are reduced in two H3.3K36M chondrocyte lines.

of H3K36me3 is correlated with the reduction of H3K36me2 (fig. S6C), which suggests that both H3K36me2 and H3K36me3 are altered to a similar degree within gene bodies of a large fraction of genes. ChIP-PCR experiments confirmed the reduction of H3K36me3 in gene bodies of four selected genes in two H3.3K36M mutant lines (fig. S6D). Finally, depletion of SETD2 and MMSET reduced levels of H3K36me3 in gene bodies, with MMSET depletion having a lesser effect than SETD2 depletion (figs. S5, B and C, and S6, E and F). These results indicate that reduced H3K36me3 levels in gene bodies of H3.3K36M mutant chondrocytes are mainly due to SETD2 inhibition, although MMSET inhibition may also contribute.

We also used ChIP-seq to analyze H3K36me2 and H3K36me3 levels in primary chondroblastomas. The H3K36me2 and H3K36me3 chromatin occupancy in gene bodies was reduced in the two chondroblastoma samples that we analyzed (fig. S7, A and B). Gene set enrichment analysis (GSEA) (*13*) indicated that genes with reduced levels of H3K36me2 and H3K36me3 in gene bodies in chondrocyte cell lines were also enriched in the corresponding gene sets from tumor samples (fig. S7, C and D). These results indicate that H3.3K36M mutant proteins have similar effects on the reduction of H3K36me2 and H3K36me3 in gene bodies in both primary chondroblastoma samples and chondrocyte cell lines containing the same H3.3K36M mutation.

We used H3.3K36M-specific antibodies (fig. S8A) and three different assays [Western blot, immunofluorescence (IF), and ChIP-seq] to determine how H3.3K36M mutant proteins associate with chromatin. The majority of the H3.3K36M mutant proteins were detected on chromatin, similarly to WT H3, as evidenced by chromatin fractionation and IF assays (fig. S8, B and C). H3.3K36M ChIP-seq identified 6162 overlapping peaks between two H3.3K36M mutant lines (Fig. 3A). H3.3K36M peaks were significantly enriched at promoters, gene bodies, and TESs ± 2 kbp ($P <$ 0.01) but exhibited significant depletion ($P <$ 0.01) at intergenic regions, as compared with randomly shuffled peaks of the same length (Fig. 3B). Two of these peaks were confirmed using ChIP-PCR (Fig. 3C). The levels of H3.3K36M mutant proteins correlated with gene expression levels (Fig. 3D). These results indicate that chromatin localization of H3.3K36M proteins is similar to that of WT H3.3 observed in other cell lines (*14*, *15*). Furthermore, the levels of H3K36M mutant proteins were higher in genes with reduced H3K36me2 or H3K36me3 levels than in those without changes in H3K36me2 and H3K36me3 (Fig. 3E and fig. S9A). Conversely, the levels of H3K36me2 and H3K36me3 were lower within H3.3K36M peak regions than in the surrounding regions (Fig. 3F and fig. S9, B and C). The inverse relationship between the levels of H3.3K36M mutant proteins and levels of H3K36me2 and H3K36me3 on chromatin lends support to the idea that the reduction of H3K36me2 and H3K36me3 is at least partly attributable to the inhibition of MMSET and SETD2 by H3.3K36M mutant proteins incorporated into the chromatin.

Gene expression analysis using RNA-seq indicated that the expression of 567 and 799 genes was elevated and reduced, respectively, in two H3.3K36M mutant cell lines compared with parental T/C28a2 cells (fig. S10A). In addition, the expression of intergenic regions with reduced levels of H3K36me2 was lower in H3.3K36M cells than in WT cells (fig. S10B). GSEA analysis indicated that genes with reduced expression in H3.3K36M mutant chondrocyte lines were significantly enriched among genes with reduced expression in chondroblastoma samples, whereas genes with increased expression were not enriched (fig. S10, C and D). A lack of correlation of genes with increased expression between H3.3K36M mutant lines and chondroblastoma samples was probably due to heterogeneity of chondroblastoma samples (table S1). The correlation of genes with reduced expression between the H3.3K36M mutant chondrocyte lines and chondroblastoma data sets suggests common mechanisms linked to H3.3K36M expression and the loss of H3K36me2 and H3K36me3. On the basis of our analysis of a limited number of gene loci, incorporation of H3.3K36M mutant proteins into nucleosomes occurs immediately prior to or simultaneously with

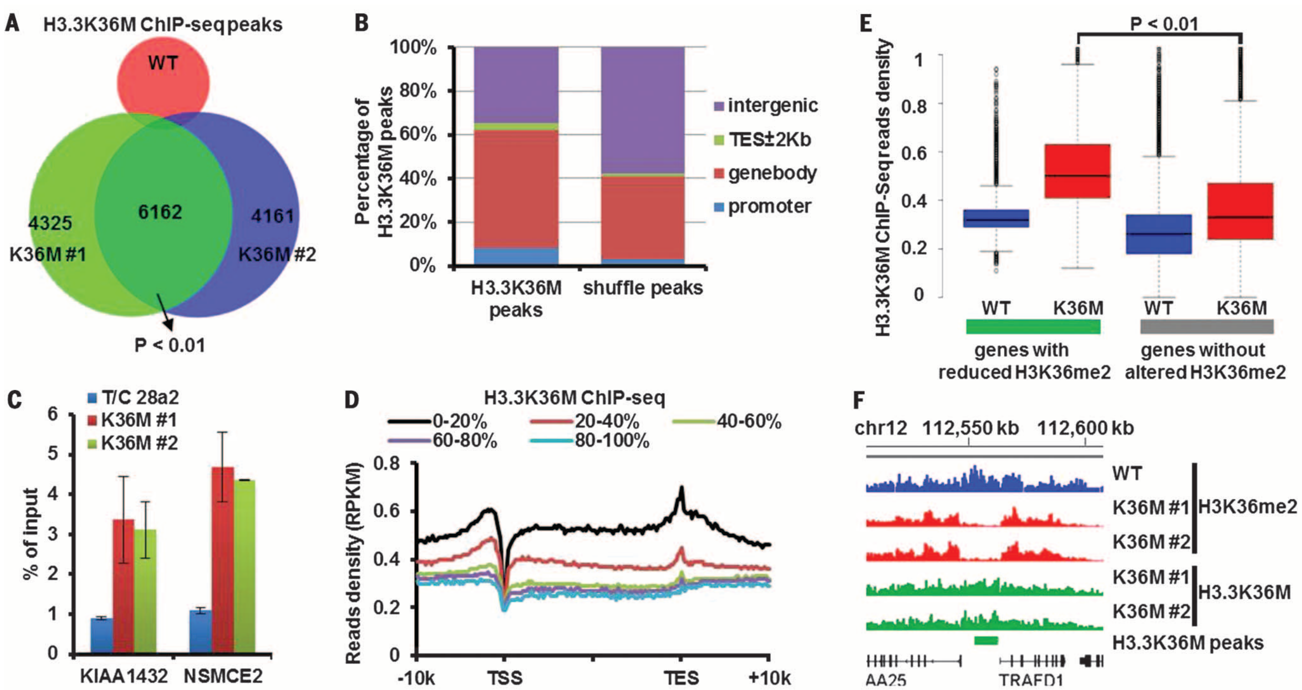


Fig. 3. H3K36M levels in chromatin are inversely correlated with levels of H3K36me2 and H3K36me3. (**A**) Venn diagram showing H3K36M ChIP-seq peaks identified in WT and two H3K36M cell lines. (**B**) Most H3.3K36M ChIP-seq peaks are in gene bodies and intergenic regions. (**C**) Validation of H3.3K36M peaks by ChIP-PCR in H3.3K36M mutant lines. Error bars indicate SD from three independent replicates. (**D**) Normalized read density of H3K36M ChIP-seq signals from 10 kb upstream of the TSS to 10 kb downstream of the TES. Genes were split into five groups, with 0 to 20% representing genes with the highest expression. RPKM, reads per kilobase per million mapped reads. (**E**) H3K36M ChIP-seq read densities in H3.3K36M cell lines for genes with or without significantly reduced H3K36me2. (**F**) Genome browser track examples of the occupancy profiles of H3K36me2 and H3.3K36M in chromosome 12 (chr12). The green bar indicates H3.3K36M peaks.

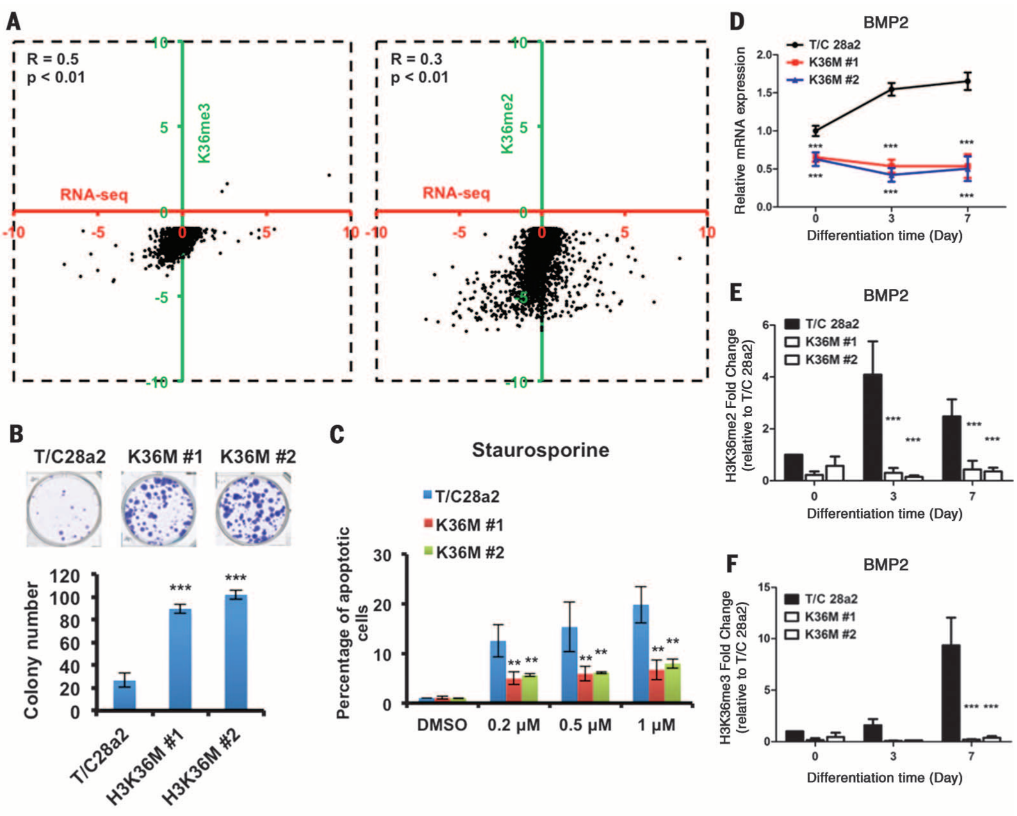


Fig. 4. H3.3K36M alters the H3K36 methylation landscape and expression of genes linked to carcinogenesis. (**A**) Log_2 fold occupancy changes in H3K36me3 (left) and H3K36me2 (right) were calculated for H3.3K36M cells versus WT cells and plotted relative to corresponding fold changes in mRNA detected by RNA-seq. Genes with significantly changed H3K36me3 or H3K36me2 in gene bodies ($P < 10^{-5}$) in both cell lines were chosen for analysis. Each dot represents a single gene. R, correlation coefficient. (**B**) H3.3K36M increases colony formation of T/C 28a2 cells. Representative images are shown at the top. (**C**) Annexin V–positive cells were analyzed and quantified by fluorescence-activated cell sorting. DMSO, dimethyl sulfoxide. (**D**) Expression of *BMP2* was analyzed by real-time PCR during differentiation conditions and normalized against glyceraldehyde-3-phosphate dehydrogenase. (**E** and **F**) H3K36me2 and H3K36me3 ChIP assays in cells differentiated for 0, 3, or 7 days. ChIP DNA was analyzed using primers amplifying the *BMP2* gene body. Fold enrichment relative to T/C 28a2 cells at the start of differentiation was calculated. Results in (B) to (F) represent the mean ± SD ($N = 3$ independent replicates, $^{**}P < 0.01$, $^{***}P < 0.001$).

the reduction of H3K36me2 and H3K36me3 and alters gene expression (fig. S11, A and B). Finally, we observed that genes with reduced levels of H3K36me2 or H3K36me3 within gene bodies exhibited a significant correlation with changes in gene expression in both cell lines (Fig. 4A) and in chondroblastoma samples (fig. S11C). These results suggest that H3K36me2 and H3K36me3 levels are associated with changes in gene expression in both cell lines and tumors. It is known that H3K36me2 and H3K36me3 antagonize H3K27me3 (*12*, *16*), raising the possibility that changes in the expression of some genes in H3.3K36M mutant cells may also be linked to deregulation of H3K27me3-repressed genes.

We performed Ingenuity Pathway Analysis (IPA) using three gene sets with altered occupancy of H3K36me2 (1143 genes) and H3K36me3 (1359 genes) and altered gene expression (598 genes) in both chondrocytes and chondroblastoma tumor samples. Genes associated with the "Molecular Mechanisms of Cancer" IPA canonical pathway were highly enriched in all three data sets (fig. S12A). Several genes assigned to the "Molecular Mechanisms of Cancer" pathway are known to be involved in DNA repair, differentiation, and apoptosis (fig. S12 and table S2). We investigated whether H3.3K36M mutant chondrocytes display cancer-associated cellular phenotypes, including DNA repair, in which H3K36me3 has a role (*17*–*19*). As determined by MTT assay, the H3.3K36M mutations did not affect proliferation of chondrocyte cell lines (fig. S13A) but increased the ability of these cells to form colonies (Fig. 4B). Moreover, H3.3K36M mutant chondrocyte cells were less sensitive to staurosporine-induced apoptosis (Fig. 4C) and formed denser micromasses when placed into chondrocyte differentiation medium (fig. S13B). The effect of H3.3K36M mutation on staurosporine-induced apoptosis could be detected at the same time when H3.3K36M mutant proteins were incorporated into chromatin (fig. S13C). Finally, the H3.3K36M mutant cells were also defective in homologous recombination (fig. S13D) but had no apparent defects in nonhomologous end joining or mismatch repair (fig. S13, E and F). Thus, H3.3K36M mutant cells exhibit several cancer-associated cellular phenotypes.

Consistent with differentiation defects, expression of *BMP2* and other genes that regulate chondrocyte differentiation (*20*) was reduced in H3.3K36M cell lines, as determined by analysis of RNA-seq results (fig. S14A). In the micromass assays, BMP (bone morphogenetic protein) signaling is required for hypertrophic chondrocyte differentiation (*21*). The mRNA levels of *BMP2* and *SOX9* were reduced in micromass cultures of T/C28a2 cells expressing the H3.3K36M mutant protein, as compared with the parental cell line (Figs. 4D and fig. S14B). Consistent with the reduction in gene expression, ChIP-PCR results showed that H3K36me2 and H3K36me3 occupancy of *BMP2* was reduced during differentiation (Fig. 4, E and F). In addition to genes involved in differentiation (*BMP2* and *RUNX2*), the expression of two genes involved in homologous recombination (*BRCA1* and *ATR*) was reduced in H3.3K36M mutant lines, as well as in MMSET- and SETD2-depleted cells (fig. S14, C and D). Thus, the cancer-associated cellular phenotypes observed in H3.3K36M mutant cells are caused, at least in part, by alterations in H3K36 methylation and expression of key genes that regulate oncogenesis and chondrogenesis.

Our results show that H3.3K36M mutant proteins are incorporated into chromatin in a manner similar to incorporation of WT H3.3 and that these mutant proteins inhibit, at minimum, MMSET and SETD2 to reduce H3K36 methylation. Moreover, H3.3K36M mutant proteins affect various H3K36 methyltransferases differently and promote chondroblastoma tumorigenesis, probably through alterations of several cancer-related processes including colony formation, apoptosis, and chondrocytic differentiation.

REFERENCES AND NOTES

1. S. A. Qasem, B. R. DeYoung, *Semin. Diagn. Pathol.* **31**, 10–20 (2014).
2. P. B. Talbert, S. Henikoff, *Nat. Rev. Mol. Cell Biol.* **11**, 264–275 (2010).
3. E. Szenker, D. Ray-Gallet, G. Almouzni, *Cell Res.* **21**, 421–434 (2011).
4. S. Behjati *et al.*, *Nat. Genet.* **45**, 1479–1482 (2013).
5. P. W. Lewis *et al.*, *Science* **340**, 857–861 (2013).
6. K. M. Chan, J. Han, D. Fang, H. Gan, Z. Zhang, *Cell Cycle* **12**, 2546–2552 (2013).
7. F. Zhang, Y. Wen, X. Guo, *Hum. Mol. Genet.* **23**, R40–R46 (2014).
8. M. B. Goldring *et al.*, *J. Clin. Invest.* **94**, 2307–2316 (1994).
9. A. J. Kuo *et al.*, *Mol. Cell* **44**, 609–620 (2011).
10. E. J. Wagner, P. B. Carpenter, *Nat. Rev. Mol. Cell Biol.* **13**, 115–126 (2012).
11. Y. Li *et al.*, *J. Biol. Chem.* **284**, 34283–34295 (2009).
12. W. Yuan *et al.*, *J. Biol. Chem.* **286**, 7983–7989 (2011).
13. A. Subramanian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
14. A. D. Goldberg *et al.*, *Cell* **140**, 678–691 (2010).
15. C. Jin *et al.*, *Nat. Genet.* **41**, 941–945 (2009).
16. G. G. Wang, L. Cai, M. P. Pasillas, M. P. Kamps, *Nat. Cell Biol.* **9**, 804–812 (2007).
17. D. K. Jha, B. D. Strahl, *Nat. Commun.* **5**, 3965 (2014).
18. S. X. Pfister *et al.*, *Cell Reports* **7**, 2006–2018 (2014).
19. F. Li *et al.*, *Cell* **153**, 590–600 (2013).
20. E. Kozhemyakina, A. B. Lassar, E. Zelzer, *Development* **142**, 817–831 (2015).
21. M. Barna, L. Niswander, *Dev. Cell* **12**, 931–941 (2007).

ACKNOWLEDGMENTS

We thank M. Goldring for T/C28a2 cells and Z. Lou for plasmids, B. Eckloff and T. Wood for sequencing, and B. Madden for mass spectrometry analysis. These studies were supported by the NIH [grants GM81838 and CA157489 (Z.Z.), AR68103 (J.J.W.), AR65397 (E.W.B.), AR049069 (A.J.v.W.), and DK58185 (T.O.)], the Natural Science Foundation of China [grant 31210103914 (R.-M.X.)], and the Epigenomics Program (Center for Individualized Medicine) of the Mayo Clinic. ChIP-seq and RNA-seq data sets have been deposited into the Gene Expression Omnibus under accession number GSE7523.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6291/1344/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S4
References (*22*–*48*)

6 December 2015; accepted 16 May 2016
Published online 26 May 2016
10.1126/science.aae0065


BOYALIFE®
For Better Health
Stem Cell Technology
Regenerative Medicine & Anti-aging
Innovative Drug Discovery
Boyalife's Core Business Activities
Cloning Technology
Disease Mechanistic Research
Cancer Immunotherapy
Genomics & Precision Medicine
ADVANCING FRONTIER IN BIOMEDICAL RESEARCH
Boyalife Group,
along with its affiliated companies,
was founded in July 2009 in China as an industrial-research
alliance among top research institutes for stem cell and
regenerative medicine.
In 2012, Boyalife Group won the prestigious
"2012 CCTV China Brand Award",
emerging as a leading brand in healthcare in China.
Start your career at Boyalife!
Contact:career@boyalife.com
www.boyalifegroup.com

How it works...

PNAS Submission & Editorial Review

Submissions are accepted from researchers all over the world.

Authors do not need to have a connection to an NAS member to publish in PNAS.

More than 75% of published papers are submitted directly to PNAS—

not contributed by NAS members.

PNAS is committed to transparency in its editorial review process.

FAQs for authors explaining the review process are available at www.pnas.org/site/authors/authorfaq.xhtml.

All articles are evaluated solely on their scientific merit by peers—

not by staff editors. Accepted papers must be of exceptional scientific importance and intelligible to a broad scientific audience.

3 TIER

PNAS uses a three-tier review process.

An NAS member makes the final decision for every paper

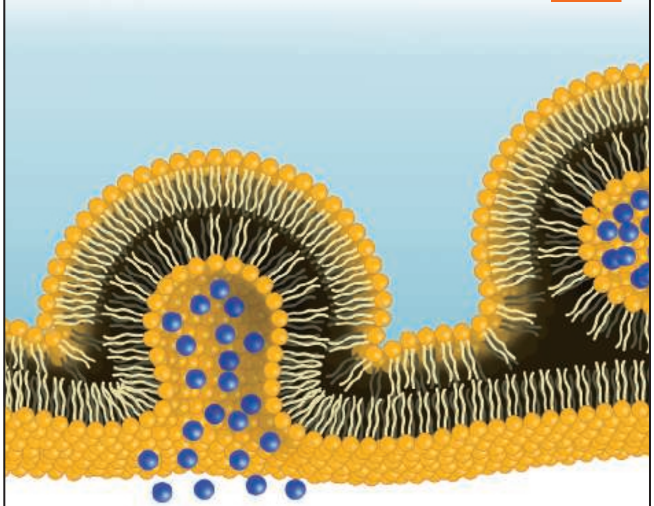
1. **Editorial Board Member—**from one of the 31 NAS disciplines. A full list of Editorial Board members and their disciplines is available at www.pnas.org/misc/masthead.xhtml.
2. **NAS Member Editor—**professional scientists and active researchers.
3. **Independent Peer Reviewers—**recognized subject experts.

PNAS

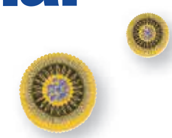
www.pnas.org

LIFE SCIENCE TECHNOLOGIES

EXOSOMES/MICROVESICLES



Membrane messengers: Extracellular vesicles



Previously dismissed as cellular debris, extracellular vesicles are currently seen as “messages in a bottle,” rife with RNA and signaling proteins. These vesicles play key roles in both development and disease. Now, researchers are developing new tools and strategies to study them and to exploit their potential in both diagnostics and therapeutics.

See the full story on page 1349.

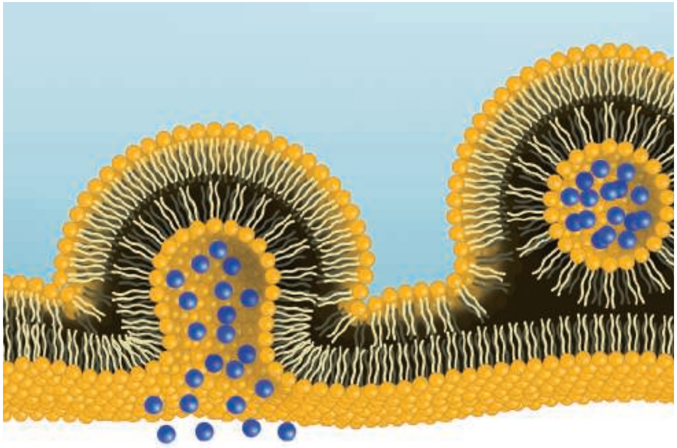
Upcoming Features

Proteomics—July 15

Genomics—October 7

Neurotechniques—November 6

Produced by the *Science*/AAAS Custom Publishing Office



Membrane messengers: Extracellular vesicles

For many years, it seems, researchers have had only a limited understanding of cellular communication. That cells could talk to one another via secreted hormones and growth factors was well known. That they also communicate using elaborate vesicular messages written in nucleic acids, proteins, and lipids was not. These vesicles play key roles in both development and disease. Now, researchers are developing new tools and strategies to study them and to exploit their potential in both diagnostics and therapeutics. **By Jeffrey M. Perkel**

In 2007, Johan Skog, a new postdoc in Xandra Breakefield's laboratory at **Harvard Medical School,** tried to culture fresh human glioblastoma tissue from biopsied material given to him by a neurosurgeon. When he put the cells into culture and looked at them under a light microscope, "they were the weirdest-looking cells we ever saw," Breakefield recalls. "They were covered with bumps. And we were thinking, 'What is this?'"

"This," as it turns out, was vesicles. Lots and lots of vesicles, some as large as half a micron in diameter. Under an electron microscope, the cells were actually pumping out much smaller particles, too—as many as 10,000 a day, Breakefield says.

Extracellular vesicles (EVs), membrane-encapsulated packages secreted by cells into the circulatory system and found in all bodily fluids, can be as large as 2 microns and as small as about 50 nm in diameter; exosomes, one particularly well-studied subtype of EV, range from 50 to 150 nm. Researchers have been aware of them for years. But until the mid-2000s, they were largely dismissed as being cellular debris or perhaps carriers of interesting protein signals. As Skog recalls, that struck him as odd. "I was looking at these vesicles and thinking, 'It would be strange if they could not contain RNA.'"

As it turns out, they did: The vesicles were chock-full of RNA that reflected the mutational status of the original tumor. Equally significant, those RNAs could serve as intracellular messengers, inducing recipient cells to change their behavior. Add purified glioblastoma vesicles to endothelial cells in culture, for instance, and they initiate blood vessel formation, or angiogenesis, says Breakefield. "They just about form tubules in front of your eyes."

"Message in a bottle"

Today, researchers like Breakefield and Skog (who after her postdoc went on to found a company dedicated to this new science, called **Exosome Diagnostics**), are working hard to tease apart the biology of EVs and to translate that information into clinical action. They have made some exciting observations. For instance, EVs' contents do not necessarily match that of the cells from which they arise.

"They're like a message in a bottle," says Andrew Hill of the Department of Biochemistry and Genetics at **La Trobe University** in Melbourne, Australia, the president-elect of the International Society for Extracellular Vesicles (ISEV).

The research community is finally taking note. Nearly 4,600 publications in PubMed include the keyword "exosomes," 4,200 of them published since 2006. In 2013, the **National Institutes of Health** launched its Common Fund-backed Extracellular RNA Communication Consortium (ERCC). According to Julie Saugstad, an associate professor in the Department of Anesthesiology and Preoperative Medicine at **Oregon Health and Science University,** who is funded under the ERCC, the ISEV annual meeting has grown from hundreds of attendees at the inaugural 2011 meeting, to thousands. "It's like everyone just woke up one day and said, 'Oh my, these are very cool.' And they are," she agrees. But that doesn't mean they're easy to study.

One method does not fit all

Esther Nolte-'t Hoen, assistant professor of Biochemistry and Cell Biology at the **University of Utrecht**, has been investigating EVs since she was a graduate student. Back then, EVs were isolated (or really, enriched) from biofluids via differential centrifugation or density-based fractionation. Those techniques are still widely used today, she notes, but they are not terribly practical for nonexperts. Moreover, they cannot be used to purify vesicle subpopulations that differ in molecular composition and function.

More recently, size-exclusion chromatography has been added to the toolbox, as have various commercial options, including the Total Exosome Isolation reagents from **Thermo Fisher Scientific,** which recover exosomes via precipitation, and the exoRNeasy Serum/Plasma Kits from Exosome Diagnostics (distributed by **Qiagen**), which rely on spin columns.

The limitation of all these methods, says Alexander "Sasha" Vlassov, Senior Manager for R&D at Thermo Fisher Scientific, is that they aren't specific for any one class of vesicles. Cells secrete "at least 10 different types of nanovesicles, but they are very difficult or impossible to differentiate due to similar size, density, and surface markers."

At least some functions attributed to vesicles may in fact be carried out by free ribonucleoprotein complexes, which also tend to copurify with vesicles. And all these different entities, whether membrane-enclosed are not, are likely formed via different pathways, carry different cargoes, and perform different functions. Exosomes, for instance, are produced via the endosomal pathway—they are made inside the cell. Some larger vesicles, in contrast, bleb from the cell surface like viruses. cont.>

Upcoming Features

Proteomics—July 15 ■ Genomics—October 7 ■ Neurotechniques—November 4

Protein power

Naturally, researchers are working to identify protein markers that can aid in distinguishing these different vesicles. Clotilde Théry, a principal investigator at the **French National Institute of Health and Medical Research (INSERM)** and the **Institut Curie** in Paris, France, for instance, recently used liquid chromatography-mass spectrometry to probe the protein content of different vesicle fractions. Her analysis identified at least six distinct vesicle classes: large, medium, and small extracellular vesicles (which can be distinguished by centrifugation), with the latter category further subdivided into four subclasses based on protein signatures, including the general vesicle markers CD9, CD63, and CD81.

Cancer cells, of course, produce unique constellations of proteins, and researchers are particularly interested in identifying proteins that mark cancer-derived vesicles. In one recent example of such research, Raghu Kalluri, professor and chairman of the Department of Cancer Biology at the **University of Texas MD Anderson Cancer Center** in Houston, and colleagues identified glypican-1 (GPC1) as a pancreatic cancer-specific vesicle marker. (Kalluri cofounded and holds equity in Codiak Biosciences, a company that exploits exosomes for the diagnosis and treatment of various diseases.) GPC1-positive vesicles seem doubly informative, Kalluri says: GPC1-positive vesicle abundance correlated with disease severity, while genetic analysis of vesicular RNA using quantitative polymerase chain reaction (qPCR) revealed a tumor's mutational status.

Similarly, researchers at **Exosome Sciences** demonstrated recently in a study of 78 former professional football players that extracellular vesicles enriched in the neurological protein tau—which the company calls "TauSomes"—are elevated in athletes with chronic traumatic encephalopathy, a neurological condition that currently cannot be definitively diagnosed antemortem.

Go it alone

Though many researchers study EV populations en masse, John Nolan, a professor at the **Scintillon Institute** in San Diego, California, prefers a different approach.

Just as cells differ in their protein and gene-expression properties, so too do their secreted vesicles. The only way to get at those differences, Nolan says, is to analyze those particles one by one. His method of choice: flow cytometry.

Adapting flow cytometry to nanometer-sized vesicles isn't easy, Nolan notes. A 10-micron cell might have 100,000 copies of an abundant protein on its surface, and even low-abundance proteins are present at a few thousand copies. But an EV measuring 100 nm in diameter is 100 times smaller than that, with 10,000 times less surface area and 1 million times less volume, and thus contains far fewer proteins for antibodies to latch onto.

Nolan and his team have built custom instrumentation designed to maximize light generation and detection, and coupled it with exceptionally bright fluorophores, protocols, and calibration standards. "We make a bunch of changes on the light source, the fluidics, the detectors, and the light collection, all of which improves performance by 50 percent to 200 percent," he explains.

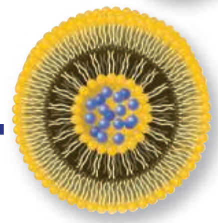
Nolan used that system to quantify vesicles studded with specific markers—CD61 and annexin V—in rat plasma. He could also distinguish particles based on size, because mixing EVs with di-8-ANEPPS, a membrane-binding dye, produces a signal proportional to the surface area of the membranes.

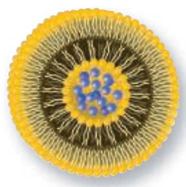
According to Nolan, commercial systems like the **Beckman Coulter** CytoFLEX are now coming online that also have the fluorescence sensitivity required for vesicle analysis. That should make the technology more accessible to the wider research community, he says. But he notes that there's still a challenge to understanding vesicles as biomarkers of disease: Nobody knows what a normal vesicle distribution looks like.

Still, researchers are forging ahead. For instance, Nolte-'t Hoen and her coworkers have used a modified **BD Biosciences** Influx system to sort vesicles derived from mast cells. Using antibodies to either CD9 or CD63, they demonstrated that some vesicles contain one protein and some contain the other. "We think it may have to do with their route of biogenesis, that they come from different parts of the cell," she says.

It's also unclear whether the two vesicle types perform different functions. And that may not be easy to determine, says Nolte-'t Hoen, "because of course now you have an antibody attached to your vesicle, which may influence the functionality." To circumvent that problem, she is now investigating "negative sorting" strategies, in which vesicles are enriched based on the proteins they do not contain.

Exosomes are produced via the endosomal pathway—they are made inside the cell.





Liquid biopsies

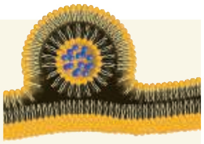
Despite the yawning knowledge gap in basic vesicle biology, many researchers' eyes are fixed elsewhere, specifically on EVs' clinical potential.

Many on the diagnostics front, for instance, are pursuing so-called "liquid biopsies." Rather than diagnosing, staging, and monitoring disease (especially cancer) via a solid tumor biopsy or noninvasive imaging, clinicians can theoretically extract similar information from blood, urine, and other biofluids such as circulating tumor cells, circulating tumor DNA—and exosomes.

Exosome Diagnostics' ExoDx *Lung(ALK)* assay, for instance, uses quantitative real-time PCR to profile the *EML4-ALK* gene fusion from blood plasma, a genetic marker of susceptibility to the kinase inhibitor crizotinib. A recent study in *JAMA Oncology* suggests that the company's prostate cancer assay, which is not yet commercially available, can likewise stratify patients into low- and high-risk categories based on the expression of three genes.

At **Caris Life Sciences**, a "molecular-profiling" company that focuses on EVs, researchers use a subtractive-binding technology called "ADAPT" (Adaptive Dynamic Artificial Poly-ligand Targeting) to identify protein signatures of disease, says David Spetzler, the company's chief scientific officer. Recently the company used a library of 2,000 peptides on 500

FEATURED PARTICIPANTS



BD Biosciences
www.bdbiosciences.com

Beckman Coulter
www.beckmancoulter.com

Caris Life Sciences
www.carislifesciences.com

Center for Systems Biology
csb.mgh.harvard.edu

Exosome Diagnostics
www.exosomedx.com

Exosome Sciences
www.exosomesciences.com

Harvard Medical School
hms.harvard.edu

Inserm
english.inserm.fr

Institut Curie
www.institut-curie.org

La Trobe University
www.latrobe.edu.au

Northwestern University
www.northwestern.edu

Oregon Health and Science University
www.ohsu.edu

Qiagen
www.qiagen.com

Scintillon Institute
www.scintillon.org

Thermo Fisher Scientific
www.thermofisher.com

University of Texas MD Anderson Cancer Center
www.mdanderson.org

University of Utrecht
www.uu.nl

patient samples to develop a signature that was better able to detect cancer in dense breast tissue than was mammography.

Similarly, Saugstad has studied the RNA content of cerebrospinal fluid (CSF) to identify a potential signature of Alzheimer's disease. Starting from a set of 756 known human microRNAs (miRNAs), her team identified 36 whose abundance in CSF appears to correlate with the disease. Given that miRNAs are regulatory noncoding molecules, that information could identify novel proteins involved in pathophysiology, she says.

In Boston, Hakho Lee, director of the Biomedical Engineering Program at the **Center for Systems Biology** at Massachusetts General Hospital, is taking yet another approach to exosome diagnostics: microfluidics.

Lee has developed liquid biopsy analytical tools based on multiple principles over the years, including electrochemical detection, magnetics, acoustics, and more. His current state-of-the-art technology, he says, exploits surface plasmon resonance (SPR).

SPR is a mature technology that has been commercialized by companies such as Biacore (now part of GE Healthcare) to quantify protein-protein and protein-ligand interactions. To generate SPR, antibodies are affixed to a thin sheet of gold atop a prism. Light passing through the prism bounces off the bottom of the gold strip at a defined angle. As molecules bind to the opposite face of the gold sensor, that angle changes in proportion to the degree of binding, providing a real-time readout of molecular interaction.

In Lee's version of the technology, antibodies are spotted on tiny gold sensors in a "periodic nanohole array," which is arranged on a microfluidic chip. As vesicles bind to these sensors, their spectral responses change proportionally to the degree of binding. Best of all, the measured vesicles can subsequently be purified for downstream genetic or protein analysis.

According to Lee, the system is highly scalable. In a proof of principle study, for instance, his lab developed a "nano-plasmonic exosome" (nPLEX) sensor with 1,089 detection sites. From a pool of 71 proteins expressed on ovarian cancer cell lines, they identified a two-protein exosomal signature that they subsequently applied to 20 cancer and 10 control subjects. That signature tracked treatment response in the nPLEX assay, Lee notes, with the marker expression dropping in responding patients but increasing in nonresponders.

Therapeutic exosomes

Researchers are also investigating exosomes as vehicles for delivering therapeutics. EVs, says Joshua Leonard, associate professor of Chemical and Biological Engineering at **Northwestern University,** seem to exhibit some of the properties—especially low toxicity—that researchers have been struggling to achieve with synthetic vesicles.

For instance, researchers can load EVs with specific cargo using electroporation, or by expressing nucleic acids in EV-producing cells. In 2011, Matthew Wood and colleagues at the University of Oxford used both approaches to show that they could use exosomes to downregulate neuronal gene expression in the mouse brain by loading the exosomes with a neuron-targeting peptide and specific short interfering RNAs. That result, Leonard says, suggests EVs can overcome at least three significant hurdles: crossing the blood-brain barrier, getting taken up specifically by neurons, and successfully delivering content inside the cells.

More recently, Leonard's team, led by graduate student Michelle Hung, has begun teasing apart the rules governing RNA-loading in EVs. The team fused an exosomal protein to a bacteriophage protein normally involved in loading nucleic acids into viruses, coupled that protein's signal sequence hairpin to RNAs of different lengths, and monitored what RNAs ended up in the resulting particles. All RNAs could be loaded, they found, though longer sequences and messenger RNAs tended to load less efficiently. "We tried to come up with a first pass at a quantitative map of the rules for loading EVs," he explains.

It will take considerable effort to convert such observations into clinical realities, of course. But given the engagement of the research community, expect those advances sooner rather than later. "There's probably a language here, and we're [only] at the level of knowing something about the alphabet," concedes Breakefield. "We don't know the grammar, we don't know who's talking to whom, or when, or why. But we're figuring it out."

Jeffrey M. Perkel is a freelance science writer based in Pocatello, Idaho.

DOI: 10.1126/science.opms.p1600106

Learn more about new methods for detecting EVs in our recent webinar
bit.ly/imageflow

HiCen GT Centrifuge
The HiCen GT Centrifuge has impressive build quality and is packed with many advanced features. The HiCen GT has a maximum speed of 14,000 rpm with a g-force of 21.913 × g. A range of fixed-angle or swing-out rotors are available. It uses the Automatic Positive Rotor Identification system of automatic rotor detection to ensure complete safety. The HiCen GT sits comfortably on any benchtop and can accommodate samples of up to 1,000 mL. The HiCen range features an assortment of models both with refrigeration and without. A touch panel and color thin-film transistor display ensure that all settings can be quickly input so that the user has complete control either manually or by selecting one of 30 programs. All components are rigorously tested and are quickly and easily changed. The HiCen GT Centrifuge operates with low noise levels because of its friction-free drive mechanisms and well-balanced rotor.
Herolab
For info: +44-(0)-1223-515440
www.herolab.de

Cell Isolation Products
Celase GMP is a proprietary enzyme containing collagenase, neutral protease, and buffer salts that are produced using avian and mammalian tissue-free raw materials, and are aseptically processed, sterile-filtered, and highly purified under good manufacturing practices (GMP) guidelines. A single, sterile, ready-to-use enzyme, Celase GMP is ideal for a wide range of cell isolation studies (adipose stem cell, biomedical, and bioprocessing) in laboratories looking to facilitate a smooth transition from bench and animal research to downstream clinical applications. Not all research applications require GMP-grade enzymes in early-phase studies. However, recent FDA guidance issued for tissue and cell products specifically cites that GMP-grade reagents should be used in drug-type validated processes. Subsequently, both regenerative medicine researchers and clinicians are now looking for GMP-quality products with a smoother regulatory approval process as the goal. The enzyme advances adipose-based research programs from preclinical to clinical levels and eliminates costly and time-consuming bridging studies.
Worthington Biochemical
For info: 800-445-9603
www.worthington-biochem.com

Field-Flow Fractionation
A series of new products for Field-Flow Fractionation (FFF) have now arrived. A new Size Exclusion Chromatography (SEC) option, together with its powerful new NovaSEC software, is now available. The new Postnova SC2000 Modular SEC option offers unrivaled flexibility in advanced SEC/gel permeation chromatography analysis. It is the first truly modular multidetector SEC system available, allowing flexible access to a wide range of applications. The economical, versatile, and scalable SC2000 is an exciting option for the Postnova FFF Characterization Platform. Offering both FFF and SEC capabilities, the 2000-Series Characterization Platform will for the first time uniquely provide laboratories a single platform able to separate both particles and molecules. Furthermore, the PN3310 Viscometer Detector perfectly matches the high-performance 21-angle Postnova Multi-Angle Light Scattering Detector and offers users a high-performance Triple Detection System for both SEC and FFF.
Postnova
For info: +44-(0)-1885-475007
www.postnova.com

Large Capacity Bioprocessing Centrifuge
The Sorvall BIOS 16 centrifuge offers an increased capacity of up to 16 L of cell culture product per run and unique design features to make working with large volumes easier and more convenient. Designed to enhance the user experience, these features include auto-door function, auto-lid function, centri-touch interface, accumulated centrifugal effect function, centri-vue remote monitoring, and eco-spin technology. The centrifuge offers the flexibility to select higher-capacity rotors or match existing workflows, with a choice of four swinging bucket rotors ranging from 6 × 1,000 mL to 8 × 2,000 mL. This product is listed with the FDA, and complies with the latest global safety standards, including UL and CE. Bioprocessing centers around the world can reap the benefits of this unique, user-friendly system designed to provide high-throughput sample processing; simple, quick setup and remote monitoring; energy savings; global safety standards; and enhanced ergonomics with every run.
Thermo Fisher Scientific
For info: 800-955-6288
www.thermofisher.com

Exosome Research Products
AMS Biotechnology (AMSBIO) has expanded its range of quick, easy, and efficient exosome isolation tools that now includes reagents, immunobeads, and immunoplates for overall or specific intact exosome isolation from small volumes of biological fluids or cell media samples. In addition, AMSBIO has introduced new efficient exosome DNA and RNA extraction kits; highly purified, lypophilized exosome calibration standards for quantitation of exosome-derived markers; ready-to-use exosome fluorescence-activated cell sorting kits for exosome isolation and profiling; highly purified exosomes for downstream analysis; validated exosomal monoclonal/polyclonal antibodies; and exosomes isolated from stem cells for cell culture. To enable precise quantification and multiple marker analysis in a number of pathological conditions including inflammation, cancer, and neurodegenerative diseases, a new range of exosome enzyme-linked immunoabsorbent assay plates is now available. Formulated to be affordable and easy to use, AMSBIO's new expanded suite will facilitate researchers' understanding of exosomes.
AMS Biotechnology
For info: +44-(0)-1235-828200
www.amsbio.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

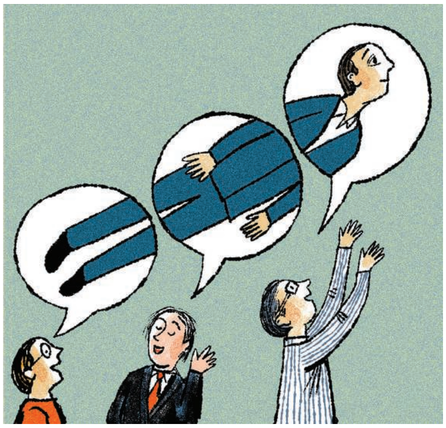
Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

By **Piotr Wasylczyk**

Three lessons rarely taught

After earning two advanced degrees, completing three postdocs, working in three countries, and finally reaching the stage when I am setting up my own lab, I realize that three lessons taught by three great mentors have influenced how I think about doing science. These lessons, each of which came at just the right time in my career, have helped me probe new intellectual territories and enjoy my work. Looking back, I appreciate the way that my mentors supported my development as a researcher and imparted valuable advice that still guides how I approach my work and career. Now, as I am moving into the role of adviser myself, I hope to be able to pass these lessons on to my current and future students.

Play around. The first lesson came from my Ph.D. supervisor in my home country of Poland. Although never explicitly voiced, he taught me that the lasers, meters, electronics, and other equipment were toys that we should play with: Open the box, tweak the knobs, and see what happens. While building laser setups, I gained hands-on experience in designing optics, mechanics, and electronics; and in soldering, welding, and machining various materials. I now appreciate his trust in my beginner's skills and the chance to learn from my mistakes.



"Three great mentors have influenced how I think about doing science."

Be sure to have fun. The next lesson came from my adviser for my first postdoc, in the United Kingdom. Having completed my Ph.D. and published some papers, I was starting to feel like a full member of the scientific community, and I was becoming confident about what I knew and was able to achieve. But I hadn't really thought deeply about why I was following the path I was on until my adviser said to me, "It is only worth doing science if you are still having fun doing it."

This notion prompted me to consider whether I was actually enjoying my research—and what I could do about it if I wasn't. With this new mindset, I gradually refocused my work over the following months and years. Instead of concentrating on planning and pursuing my next career steps and applying for grants, I spent more time exploring new research areas and intellectually wandering in search of attractive ideas. I dabbled in experimental photonics and microfabrication and ended up in smart materials and robotics. Changing research topics is always risky, but as I transitioned between disciplines, I discovered that novelty generates a wave of excitement and that gaining new perspectives unleashes great intellectual potential.

Find what suits you. The third lesson came on my first day in a new postdoc, this time in Italy. My boss, who directed a few research teams, told me, "Look around at all the groups. See what they do, and find what suits you best. I believe in self-organization; everyone should do what they like, as only then will they do it with joy and passion." I was surprised to be offered this level of independence, but I followed his advice and ended up embarking on an ambitious project—designing and building bacteria-sized robots powered and controlled by light—that was quite far from my initial plans. I wondered what it would mean for my career if we failed, but I felt that this was the area I could explore, as my adviser advocated, "with joy and passion."

Talking to other scientists, both young and mature, I see how difficult it can be to enjoy research. I feel privileged that my mentors encouraged me to play, have fun, and pursue joy, and that their support afforded me the opportunity to take risks. As for the robot, after 2 years of work, our team built a machine half the size of the world's smallest known insect, a male parasitic wasp. The success was gratifying, but I think I would have been happy with my decision to push my limits regardless of the results.

Now, as I put the final touches on the design of my new lab space, I plan to have a poster at the entrance to help remind my group—and myself—that taking risks is the essence of research. It says, "Which research project would you start today if you were certain you would succeed?" ■

Piotr Wasylczyk is an assistant professor in the Faculty of Physics at the University of Warsaw. Send your story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

JOB FOCUS: MICROBIOLOGY

IMMUNOLOGY FACUTY POSITION

The Department of Immunology and Microbiology in the Wayne State University School of Medicine seeks outstanding tenure track **ASSISTANT AND ASSOCIATE PROFESSOR** candidates who are pursuing fundamental immunological problems in areas that include inflammation-related disease, mucosal immunity, neuroimmunology, cancer immunobiology, and immunobiology and immunologic control of commensal and pathogenic microbes. Candidates will be expected to establish a high-impact extramurally funded research program, and to participate in teaching and service. A competitive start-up package commensurate with the candidate's experience and achievement will be provided.

For further information about this position, and to submit your application, please visit our online application site at website: **https://jobs.wayne.edu (Posting 041955)**.

Wayne State University is an Affirmative Action/Equal Employment Opportunity Employer, and is committed to nondiscrimination and equal opportunity for all persons.

VIB is a non-profit life sciences research institute located in Flanders, Belgium. About 1,450 scientists and technicians conduct basic research on the molecular mechanisms that are responsible for the functioning of the human body, plants, and microorganisms. Through a close partnership with all Flemish universities and a solid funding program, VIB unites the forces of 75 research groups clustered in 8 research centers. The corner stone of VIB's policy is excellence, as well in science as in technology transfer. The goal of the research is to move forward the boundaries of our knowledge of life profoundly. Through its technology transfer activities, VIB wants to convert research results into products for the benefit of consumers and patients. VIB develops and disseminates a wide range of scientifically substantiated information about biotechnology.

The University of Antwerp is a young, dynamic and forward-thinking Belgian university. Every day, more than 5000 employees' efforts ensure that innovative education is provided to over 20.000 students, that ground-breaking fundamental and applied research is conducted and that the university is able to play its important role as a service provider in society. Striving for international excellence, the University of Antwerp is a true research university, with a particularly strong expertise in the following key areas of research: Imaging; Neurosciences; Infectious Diseases; Drug Discovery and Development; Materials Characterization; Ecology and Sustainable Development; Harbor, Transport and Logistics; Urban History and Contemporary Urban Policy; Social Economic Policy and Organization. The University of Antwerp ranks 11th worldwide among young universities.

The University of Antwerp and VIB are currently seeking outstanding candidates (m/f) to fill the position of

DIRECTOR/RESEARCH PROFESSOR

for the VIB-Department of Molecular Genetics (DMG) at the University of Antwerp

The VIB-UAntwerp Department of Molecular Genetics is building on a longstanding tradition of integrating expertise in molecular genetics and clinical neurology to analyze aging-related neurological diseases of the central and peripheral nervous system, including Alzheimer disease, frontotemporal lobar degeneration, and peripheral sensory and motoric neuropathies. The overall objectives are to enhance our current understanding of the biology of the nervous system during healthy aging and of the molecular etiology of aging-related neurological diseases. New findings will be translated to the clinic as well as to experimental work. The key goal is to gain a holistic view of the disease processes and to create effective translational pipelines and potential therapies that are specifically tailored to well-characterized patient groups.

Job description

VIB and UAntwerp are in search for a dynamic, internationally recognized leader to inspire and lead the research center. The selected candidate will be appointed as scientific director at VIB and will have a research professor appointment (ZAPBOF) as (full) professor at the University of Antwerp.

The director/research professor will provide vision and leadership to the research center, with responsibility for:
• developing the scientific strategy of the research center and its future mission
• creating a stimulating environment and dialectic culture, which fosters talent and triggers excellence
• organizing and managing the research center in terms of science, tech transfer, logistics, finance and human talent

It is expected that the director/research professor will maintain active interest in leading an independent research lab (for which significant long term support will be provided). The director/research professor will be a member of the VIB management committee in which he/she has co-responsibility for the overall success of VIB.

Lecturing may become part of the job description in the long(er) term.

Requirements

The successful candidates for the position of director/research professor:
• have a PhD degree and are expected to be experienced visionary scientists, widely recognized in their academic field, who have demonstrated a strong record of scientific publications in leading scientific journals in the field. They will continue their own research program at VIB/UAntwerp.
• rank as (full) professor with - preferentially - a track record of management in universities or research institutions.
• have an extensive international network with a wide scope of research collaborations.
• have excellent communication and negotiating skills and have a strong will for developing a common vision and purpose for VIB and their own research center.
• by preference, have experience with technology transfer and science policy.

Our offer

The total VIB/UAntwerp package consists of:
(a) an appointment as (full) professor in a tenured position. In the case of a first appointment at UAntwerp, the university board may proceed to a temporary appointment for a period not exceeding three years. After a favorable performance assessment in the position of research professor the position will become tenured.
(b) personal: • a competitive salary commensurate with experience • total social security, including pension scheme and health care program
(c) own research group: • start-up package • funding of your own research group with 5-10 positions • access to the VIB core facilities: (deep)sequencing, transcriptomics, proteomics, antibody and protein production, bio-imaging, NMR, compound screening, bio-informatics training facility (see www.vib.be)
(d) research center grant: • solid funding of the research center, subject to excellent performance with quinquennial reviews.

The date of appointment will be as soon as possible.

How to apply

• Candidates who are interested in this position are asked to send a complete CV, publication list, vision text (max. 5000 words) and 3 letters of reference to marijke.lein@vib.be. **Closing date for applications is 30 September 2016**.
• Further information about the profile and the description of duties can be obtained from Jo Bury, managing director VIB (+32 9 244 6611) or Prof. Dr. Jean-Pierre Timmermans, President of the UAntwerp Research Council (+32 3 265 3002).
• Further information regarding the terms of appointment can be obtained from Marijke Lein - HR director of VIB (+32 9 244 6611) or Greet Dielis – HR manager of the University of Antwerp (+32 3 265 3153).
• Further information on www.vib.be, www.molgen.vib-ua.be and www.uantwerpen.be

FRED HUTCH
CURES START HERE™

Senior Faculty Leadership Position in Immunology at Fred Hutchinson Cancer Research Center

The Vaccine and Infectious Disease Division (VIDD) of the Fred Hutchinson Cancer Research Center seeks exceptional applicants for a full-time senior faculty leadership position in immunology at the Full Member rank (comparable to Professor). The primary responsibility of this position will be to develop and lead a comprehensive, cross-disciplinary integrated research center (IRC) involving multiple investigators that will focus on pathogen-induced cancers. The candidate will be expected to conduct laboratory-based translational immunology research as part of this IRC and within VIDD, and an emphasis on mechanisms of memory/effector cell induction and immune dysfunction in the context of cancer or cancer-associated infections is highly desirable.

Candidates for this position must have a well established, robust, funded program that is nationally and internationally recognized for excellence in immunology, immunotherapeutic design, or viral oncogenesis.

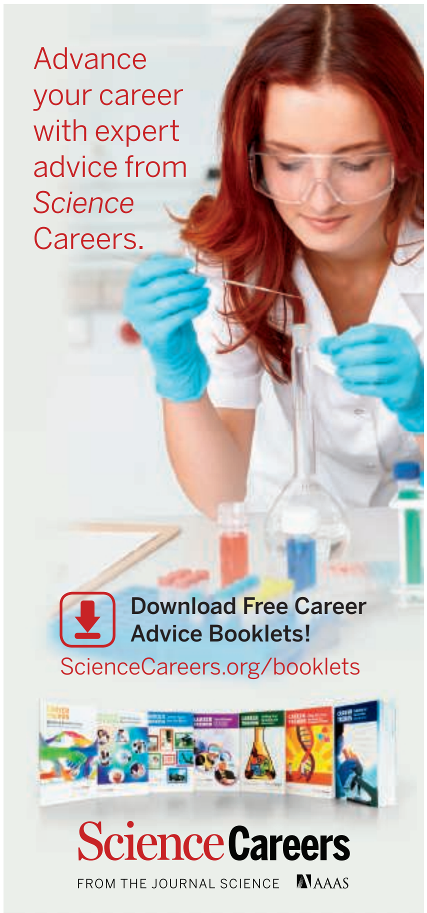
Applicants must have an MD (or foreign equivalent) or PhD (or foreign equivalent). Selection criteria include excellence in scholarship, creativity in research, and demonstrated leadership in the profession.

VIDD scientists integrate clinical care, computational methods, and basic science research in immunology, virology, and vaccine design to reduce the global burden of infectious disease.

The Fred Hutchinson offers a vibrant intellectual environment within a beautiful, lakeside campus in Seattle's South Lake Union biotech hub. VIDD occupies a new building that is connected by walking trails to Seattle Cancer Care Alliance and the other four Divisions of the Fred Hutch and by trolley to major research partners such as the University of Washington School of Medicine, Seattle Children's Research Institute, Center for Infectious Disease Research (formerly Seattle Biomedical Research Institute), and the Infectious Disease Research Institute.

Interested candidates should submit a CV, a concise research plan statement, and the names and contact information for three (3) references to: **fredhutch.org/job/6653**. Specific inquiries can be directed to **Julie McElrath** at **206-667-1858**. Applications should be received by **August 15, 2016** to assure consideration and will be evaluated as received.

The Fred Hutchinson Cancer Research Center is an Affirmative Action, Equal Opportunity Employer. All qualified applicants will receive consideration for employment without regard to, among other things, race, religion, color, national origin, sex, age, status as protected veterans, or status as qualified individuals with disabilities. We strongly encourage applications from women, minorities, individuals with disabilities and covered veterans.



NANYANG TECHNOLOGICAL UNIVERSITY | NUS National University of Singapore | SUTD SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

TEMASEK RESEARCH FELLOWSHIP (TRF)

A globally connected cosmopolitan city, Singapore provides a supportive environment for a vibrant research culture. Its universities Nanyang Technological University (NTU), National University of Singapore (NUS) and Singapore University of Technology and Design (SUTD) invite outstanding young researchers to apply for the prestigious TRF awards.

Under the TRF scheme, selected young researchers with a PhD degree have an opportunity to conduct and lead defence-related research. It offers:

- A 3-year research grant of up to S$1 million commensurate with the scope of work, with an option to extend for another 3 years
- Postdoctoral or tenure-track appointment (eligibility for tenure-track will be determined by the university)
- Attractive and competitive remuneration

Fellows may lead, conduct research and publish in these areas:

- Adaptive Camouflage Techniques and Technologies
- Cyber Security for Autonomous Systems
- Perception under Adverse Conditions for UGV Navigation

For more information and application procedure, please visit:

NTU – http://www3.ntu.edu.sg/trf/index_trf.html
NUS – http://www.nus.edu.sg/dpr/InfoForResearchers/trf.html
SUTD – http://temasek-labs.sutd.edu.sg/funding-opportunities/temasek-research-fellowship-trf

Closing date: 23 September 2016 (Friday)

Shortlisted candidates will be invited to Singapore to present their research plans, meet local researchers and identify potential collaborators in February/March 2017.

UNIVERSITY of VIRGINIA HEALTH SYSTEM

TENURED/TENURE-TRACK FACULTY POSITIONS (OPEN RANK)

The newly expanded **Center for Membrane and Cell Physiology** at the University of Virginia invites applications for tenured/tenure-track positions in High-Resolution Live-Cell and Tissue Imaging. Live-cell and super-resolution imaging are undergoing a revolution and the University of Virginia seeks to position itself at the forefront of these developments by building a team of creative and highly collaborative scientists developing and employing such methods to solve important biomedical problems. Tenure status and rank of the positions will be dependent on qualifications. Incumbents will be resident members of the Center for Membrane and Cell Physiology and will also have an appointment in a basic science or clinical department of the UVa School of Medicine. Outstanding opportunities exist to collaborate with structural, computational, cardiovascular, cancer, developmental, cell, and chemical biologists and neuroscientists in a highly interactive research environment at the University of Virginia. Competitive start-up packages will be offered.

Successful applicants will have a Ph.D., M.D. or equivalent degree, will be highly creative, and must have demonstrated exceptional scholarly success in their field. Demonstration of sustained grant or equivalent support is required for appointments at a mid-career or senior rank.

To apply, visit **https://jobs.virginia.edu** and search on **Posting Number 0618868**. Complete a Candidate Profile online, attach a cover letter, curriculum vitae, statement of research interests and contact information for three references. Review of positions will begin on **July 5, 2016**. The positions will remain open until filled.

For further information about the positions, the scope of the search and application process, please contact **Ms. Jennifer Nickerson** at **jcn6f@virginia.edu** or **Dr. Jochen Zimmer** at **jz3x@virginia.edu**.

The University of Virginia is an Equal Opportunity/Affirmative Action Employer. Women, minorities, veterans and persons with disabilities are encouraged to apply.